TEC-0110

Research in the Automated Analysis of Remotely Sensed Imagery: 1994-1995

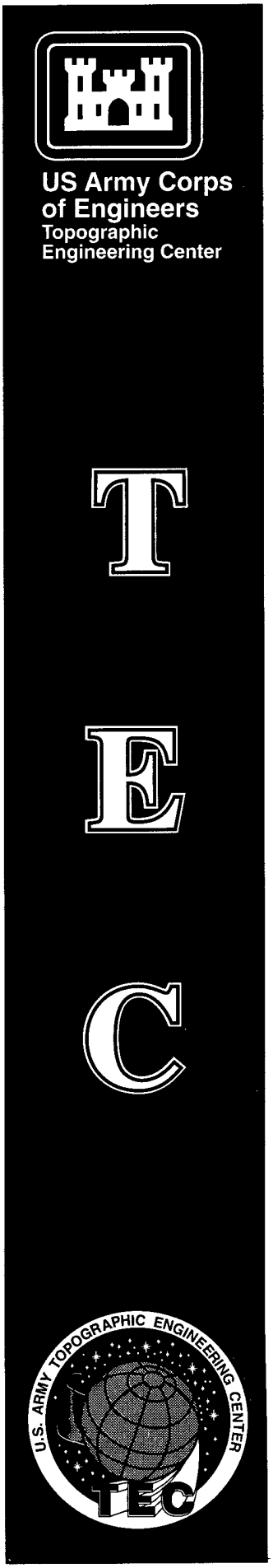
David M. McKeown, Jr., et al.

Carnegie Mellon University
School of Computer Science
Pittsburgh, PA 15213-3890

August 1998

Prepared for:
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714

Monitored by:
U.S. Army Corps of Engineers
Topographic Engineering Center
7701 Telegraph Road
Alexandria, VA 22315-3864



Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1998	Technical Sept. 1994 - Sept. 1995

4. TITLE AND SUBTITLE

Research in the Automated Analysis of Remotely Sensed Imagery: 1994-1995

5. FUNDING NUMBERS

DACA76-92-C-0036

6. AUTHOR(S)

David M. McKeowan, et al.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Carnegie Mellon University
School of Computer Science
Digital Mapping Laboratory
Pittsburg, PA 15213-3890

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
3701 North Fairfax Drive, Arlington, VA 22203-1714

U.S. Army Topographic Engineering Center
7701 Telegraph Road, Alexandria, VA 22315-3864

19. SPONSORING / MONITORING AGENCY REPORT NUMBER

TEC-0110

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report presents an overview of the program of research at the Carnegie Mellon University's Digital Mapping Laboratory in the analysis of remotely sensed imagery and the construction of virtual world databases. Progress is reported in the areas of digital photogrammetry, automatic and semi-automatic building extraction, stereo analysis, multispectral image analysis, and virtual world construction for distributed simulation.

14. SUBJECT TERMS

Site Modeling, Feature Extraction, Information Fusion, Database Intensification, Hyperspectral Image Analysis

15. NUMBER OF PAGES

48

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

This work was sponsored by the Defense Advanced Research Projects Agency (DARPA) and was monitored by the U.S. Army Topographic Engineering Center (TEC) under contract DACA76-92-C-0036. The DARPA Program Manager was Mr. Oscar Firschein, and the TEC Contracting Officer's Representative was Mr. Thomas Hay.

ACKNOWLEDGMENTS

The research presented in this report was supported by DARPA and TEC under contracts DACA76-91-C-0014, DACA76-92-C-0036, DACA76-95-C-0009, and the Air Force Office of Scientific Research (AFOSR) under Contract F49620-92-J-0318. Student research and training support was provided by the U.S. Army Research Office ASSERT program under contract DAAH04-93-G-0092.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Topographic Engineering Center, the U.S. Army Research Office, the Air Force Office of Scientific Research, the Advanced Research Projects Agency, or the United States Government.

We thank the following members of the Digital Mapping Laboratory for their contributions to various aspects of the research reported: Michael Bowling, Michael Higgins, Christian Hoffman, Nathan Segerlind, Aaron Wald, and Erik Witte.

This report appeared in the Proceedings of the 1996 DARPA Image Understanding Workshop, 12-15 February 1996, Palm Springs, CA.

For additional information regarding our research activities, please refer to the following world wide web address: http://www.cs.cmu.edu/~MAPSLab/

RESEARCH IN THE AUTOMATED ANALYSIS OF REMOTELY SENSED IMAGERY: 1994–1995

David M. McKeown, G. Edward Bulwinkle, Steven D. Cochran
Stephen J. Ford, Stephen J. Gifford, Wilson A. Harvey
Yuan C. Hsieh, Chris McGlone, Jeff McMahill
Michael F. Polis, Michel Roux*, Jefferey A. Shufelt

Digital Mapping Laboratory
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213-3890

1 Introduction

The ongoing research focus of the Digital Mapping Laboratory is the extraction of cartographic features from aerial imagery. This has proven to be a difficult problem; although progress has been made, there are still no automated systems that exhibit robust and accurate behavior across a wide range of imagery and scene complexity. Some thoughts on why progress has been slow can be found in the paper [McKeown, 1994]. Our program of research examines a variety of input data sources, including panchromatic and multispectral imagery, either in single image analysis or by combining analysis across multiple images. We attempt to employ all available image, scene, and domain knowledge in our techniques; our recent emphasis of rigorous photogrammetry is an example of this. Our algorithmic approach continues to leverage the strong points of several types of methods by fusing their results, in an attempt to combine their strengths and overcome weaknesses. A recent successful example has been the incorporation of many of our automated feature extraction and matching techniques into a semi-automated system (Section 4), combining the operator's top-level cueing and editing with the feature extraction and matching techniques used in our automated systems. A relatively new area of research is the generation of highly complex terrain surface representations to support the construction of virtual worlds for distributed simulation applications. In order to support the intensification of standard digital products from the Defense Mapping Agency (DMA) with timely and specialized information from imagery, we are investigating issues in topology and efficiency of representations of merged cartographic datasets derived from a variety of different sources.

This report presents an overview of the spectrum of current research within the Digital Mapping Laboratory and serves as a pointer to more detailed technology articles in this Image Understanding Workshop Proceedings as well as to other publications by members of our research group.

Building extraction has been, and continues to be, a major research interest of our laboratory. This report describes three separate but related approaches to the problem. Section 2 describes the development status of PIVOT, a system that extracts buildings from single images using vanishing point geometry and a set of volumetric primitives; the vanishing point analysis techniques employed by PIVOT are described in another report [Shufelt, 1996]. This is followed by a brief performance analysis of the the MULTIVIEW system (Section 3), which detects buildings using feature matching in multiple images. Finally, a semi-automated site modeling system, SiteCity, is discussed in Section 4 and is more completely described in [Hsieh, 1996].

A rigorous photogrammetric capability provides an infrastructure for the systems built in our laboratory. Recent work, described in Section 5, has studied the use of object-space geometric constraints within

*Current address: Departement Images, Telecom Paris, 46 rue Barrault 75013 Paris, France.

a simultaneous multiple-image bundle adjustment. This allows more precise modeling of buildings, and may aid in the evaluation and editing of geometric hypotheses.

Our previous work in stereo has concerned the fusion of results from area- and feature-based matchers. We are currently evaluating an iterative approach to stereo matching, developed at the U.S. Army Topographic Engineering Center (TEC) by Raye Norvelle; some comparisons between it and our existing stereo systems are given in Section 7. The fusion of stereo elevation information and building hypotheses from monocular methods also are presented.

While CMU was among the first to investigate the use of multiple cooperative methods for road detection and tracking in high-resolution aerial imagery, this avenue of research has largely been dormant until recently. Section 6 describes new modifications to the existing system, including its conversion to work in object space instead of image space and a new user interface to aid in diagnostics and evaluation.

Previous research in information fusion using multispectral imagery has involved 11 channel Daedelus scanner data at an 8 meter ground sample distance (GSD). While promising results have been obtained, we have long believed that finer spatial resolution combined with a more extensive spectral capability was required to fully exploit structural information in panchromatic imagery below a 1 meter GSD. Toward a demonstration of high-resolution land cover classification in urban areas, we have participated in the acquisition of a hyperspectral data set over Fort Hood, TX. The acquisition procedure and our processing plans are described in Section 8.

The generation of virtual worlds for training and simulation has become an increasingly important demonstration and application area for image understanding and related research. Section 9 and the companion paper [McKeown *et al.*, 1996] give an overview of our research progress in the construction of bare earth terrain models and the integration of man-made and natural features such as roads, rivers, lakes, and bridges into the simulation terrain skin. This application area has a strong relationship to our ongoing research in automated site modeling in that it relies on accurate geo-position of man-made structures, the integration of cartographic features and digital elevation models, and the generation of realistic models of actual areas of the world. One difference is that the scale of distributed simulation databases often encompass much larger areas of interest, up to several hundred kilometers on a side, with variable levels of detail and fidelity contained therein.

Finally, the development and evaluation of end-to-end image understanding systems and their underlying algorithms requires the availability of large sets of test imagery with varied characteristics. The generation of realistic simulation databases requires the fusion of digital feature and elevation data from a variety of sources. In Section 10 we describe the tools we have developed to import and access large databases of imagery and digital data.

2 PIVOT—Automated Building Extraction from Monocular Views

Building extraction from aerial images has been a topic of great interest in the computer vision community for several years. In previous work, we explored the integration of photogrammetric constraints and rigorous camera modeling into an existing monocular building extraction system, BUILD [McKeown, 1990]. The resulting system, VHBUILD [McGlone and Shufelt, 1994a; McGlone and Shufelt, 1993], exhibited two new developments in the field; the ability to generate 3-D object space models solely from monocular imagery, including peaked-roof buildings, and the ability to operate on oblique imagery. The performance of VHBUILD was qualitatively and quantitatively shown to be superior to its predecessor, validating the hypothesis that photogrammetric knowledge can serve a key role in augmenting the performance of feature extraction systems [McKeown and McGlone, 1993].

The lessons learned from these research systems, our other building extraction systems [Irvin and McKeown, 1989; Shufelt and McKeown, 1993], and the work carried out by other research groups [Nicolin and Gabler, 1987; Huertas and Nevatia, 1988; Mohan and Nevatia, 1989; Liow and Pavlidis, 1990; Fua and Hanson, 1991; Lin *et al.*, 1994; Jaynes *et al.*, 1994], suggest a set of approaches for developing a building extraction system that can achieve consistently robust performance over a wide variety of aerial imagery. In this section, we briefly describe these principles and their motivations, the implementation status of a building extraction system based on these principles, and new results in vanishing point detection that have come out of this work.

2.1 Current Work on Building Extraction

Much of the previous work in building detection and delineation relies heavily on a variety of assumptions about the scene, the buildings in it, or the imaging process. Systems operating under these constraints are able to achieve reasonable performance on the limited class of imagery that obeys these assumptions. Unfortunately, such systems typically fail on imagery outside of the specific domain for which they were designed, and it is often unclear how the assumptions for these systems can be removed in a principled way to achieve improved performance on a wider set of imagery, without sacrificing performance in the original domain.

We briefly describe here a set of ideas motivating the design of **P**erspective **I**nterpretation of **V**anishing points for **O**bjects in **T**hree dimensions (PIVOT), a single-image building extraction system currently under development at CMU. The goal of PIVOT is to achieve robust performance on images with a wide range of viewing acquisition angles, object shape complexities, object densities, shadow and object occlusions, and shadow and illumination effects, given only image acquisition parameters and date/time information.

A key design goal of PIVOT was to *rigorously model the image acquisition process and exploit the resulting geometry.* Many systems use constraints on the scene and/or the image to constrain the search space for building hypotheses. Examples include the assumption that images are acquired from a nadir viewing angle, or that buildings have a specific shape. While these constraints give a system leverage for attacking the 2-D to 3-D mapping, they also limit the generality of the system. By carefully modeling the image acquisition process via tools from photogrammetry, geometric constraints can be derived that make no restrictions on the scene or the image. Vanishing points are one such example, discussed further in Section 2.2.

To handle a wide variety of building shapes and sizes, PIVOT *uses primitives for generic shape representation.* Recent psychological theories [Biederman, 1985; Lowe, 1985; Pentland, 1986] suggest that the complexity of our visual surroundings is the result of a combination of a few basic volumetric forms, or *primitives.* PIVOT uses two simple shapes: a rectangular volume and a triangular prism, as building blocks for constructing complex buildings. Vanishing point labelings of line segments in the image are used to hypothesize the presence of the appropriate primitive.

PIVOT searches for plausible instances of primitives in an image by *using geometric constraints as soon as possible, but no sooner.* On one hand, it is necessary to use constraints to limit the hypothesis search space as early as possible to prevent combinatorial explosion; on the other, constraining interpretations of low-level edge and line data prematurely can prevent a system from correctly hypothesizing the underlying shape. PIVOT uses vanishing points to constrain the geometry of each intermediate representation, from lines to corners to chains of corners to primitives.

The primitive generation phase of PIVOT, based on vanishing point analysis, is complete; work is now focused on the construction of object space building models by a combination of primitives, and methods for verifying these object space models by illumination and shadow analysis. The vanishing point analysis

techniques used in PIVOT represent a significant improvement over previous methods for vanishing point detection, and are described in the next section.

2.2 Vanishing Point Detection in Aerial Imagery

Under a central projection camera model, a set of parallel lines in a scene projects to a set of lines in the image that converge on a single point, known as a *vanishing point*. Each vanishing point corresponds to a unique orientation in 3–space. This correspondence provides a useful technique for inferring 3-D structure from a 2-D image. Image lines that pass through a vanishing point can be assigned the corresponding orientation in object space, information that can be exploited by 3-D object detection algorithms. This approach was employed in the VHBUILD system to locate horizontal and vertical line segments in object space, that were then used to generate geometrically consistent building hypotheses. This approach also is employed in the PIVOT system currently under development.

However, existing approaches in the literature have drawbacks that limit their usefulness on the wide range of imagery over which PIVOT is intended to operate. Most techniques for detecting vanishing points use a variant of the Gaussian sphere histogramming approach developed by Barnard [Barnard, 1983], which is essentially a Hough transform on azimuth-elevation parameter space on the sphere. Since this approach uses an accumulator array in which all edges cast votes for vanishing points, no distinction is made between building edges and texture edges caused by natural patterns, which leads to noise in the array and false vanishing point solutions. Another difficulty is that aerial imagery often exhibits weak perspective effects, which leads to inaccuracies in vanishing point solutions.

To achieve robust performance on a wide variety of aerial imagery, two novel techniques were developed ugment the standard Gaussian sphere approach for detecting vanishing points. The first technique, *nitive-based vanishing point detection*, uses knowledge about the shapes of objects of interest to guide e search for vanishing points on the sphere. For example, a rectangular building sitting on the ground is comprised of lines of three orientations; a vertical orientation for the edges joining walls together, and thogonal horizontal orientations for the edges bounding the roof and floor of the building. With this knowledge, the sphere can be searched for orthogonal maxima along the great circle whose planar normal corresponds to the vertical orientation. The idea extends naturally to other polyhedral shapes.

The second technique, *edge error modeling*, models the uncertainty in position and orientation of line nts represented in discrete digital geometry. Natural terrain and vegetation in aerial imagery often produce short, randomly scattered line segments during edge detection, unlike man-made structures, that generally produce long, straight, and regular line segments. By taking these differences into account, the histogramming process becomes less noisy, since the contributions of randomly-oriented texture edges are scattered more widely over the accumulator array, reducing the likelihood of false maxima.

Together, these techniques lead to robust performance on imagery for which the standard methods are susceptible to failure. Figure 1 shows the results of edge detection and recursive line fitting on a portion of a test image, FLAT_L, distributed as part of an ISPRS test on image understanding [Fritsch *et al.*, 1994]. In this image, the grassy areas and trees contribute a significant number of randomly oriented texture edges. Figure 2 shows the line segments that pass through the orthogonal horizontal vanishing points detected by the new vanishing point analysis techniques, where each of the two horizontal orientations is colored differently. In this case, the horizontals are properly labeled despite the presence of large amounts of noise.

The vanishing points for slanted peak edges on symmetric peaked-roof buildings can be automatically detected as well, using the primitive-based detection approach described earlier. Figure 3 shows the slanted peak edges lying in the same vertical plane in object space as the dark horizontals in Figure 2; Figure 4 shows the slanted peak edges lying in the same vertical plane as the light horizontals. As we expect, many

Figure 1. Line segments for FLAT_L.

Figure 2. Horizontal line segments.

Figure 3. First pair of slanted segments.

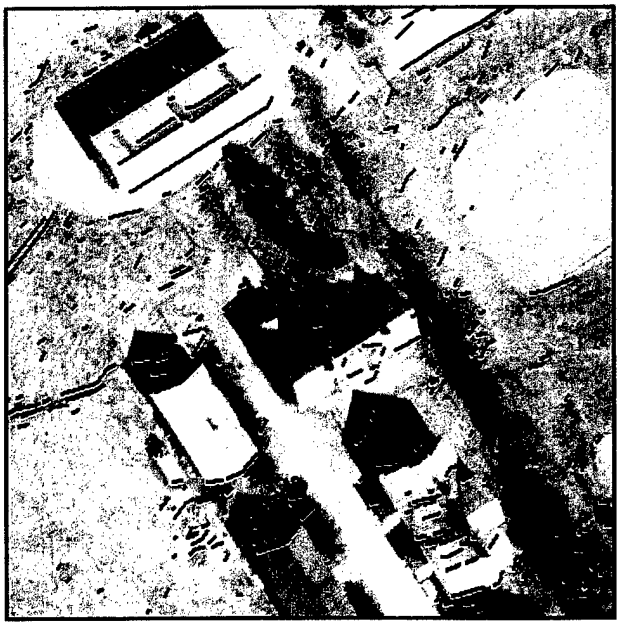


Figure 4. Second pair of slanted segments.

texture edges pass through the vanishing points in all of the figures, hence get labeled as horizontals or slanted lines. However, the important observation is that nearly all line segments corresponding to building edges have been given proper orientation labelings. It also is important to remember that these labelings correspond to 3-D orientations; the slanted edges in Figures 3 and 4 are hypothesized to form angles of 39° and 47°, respectively, with the ground plane. This illustrates the ability to infer 3-D orientations from a single view using vanishing points.

The vanishing point work is described in detail in a companion paper to this overview [Shufelt, 1996], which presents two distinct edge error models, the primitive-based vanishing point detection technique, and a thorough quantitative evaluation of performance on several aerial images, contrasting these new techniques with existing methods from the literature. These techniques are used in PIVOT to generate 3-D orientation hypotheses and constrain the search for primitives. Current research and implementation efforts are focused on the development of object space hypothesis generation and verification methods.

3 MULTIVIEW—Multiple-Image Building Extraction

In previous papers [Roux and McKeown, 1994a; Roux and McKeown, 1994b; Roux *et al.*, 1995] we presented the MULTIVIEW system for the detection and delineation of buildings by direct matching of cues in multiple images. We have performed extensive evaluations on the system in order to evaluate its performance, especially the effects of image processing order on the results.

3.1 System design overview

In MULTIVIEW we construct 3-D roof surfaces by matching salient building features extracted from different views to provide 3-D building surface cues. No assumptions are made concerning the 3-D structure of the building roof, since both flat roof and peaked roof buildings are handled identically. We begin with sparse features (building corners) extracted from multiple views of the scene. This technique relies heavily on knowledge about the imaging geometry and acquisition parameters to provide rigorous geometric constraints for the matching process.

We have been working on two different strategies: pairwise matching and multiple view integration. Pairwise matching uses two views to perform the initial construction of 3-D corners and line segments. Multiple view integration incorporates information from several views in order to solve hidden surface problems and to provide more accurate positioning of the 3-D object models.

We use epipolar, height, and orientation constraints for matching VHBUILD corners between pairwise views. These constraints are deduced from the collinearity equations, which describe the projective transformation between the camera coordinate system and the world coordinate system. These corner matchings provide 3-D corners in a local object space coordinate system. A graph is constructed in this local coordinate system where nodes represent the 3-D corners. Links between these nodes are created according to the image intensity gradient between the corners in both images.

In order to form polygonal surfaces we perform a search for cycles of corners and edges using geometric constraints such as planarity and perpendicularity in object space. To reduce the complexity of the cycle generation algorithm, only the best links, according to the image intensity gradient, are used to generate cycle hypotheses. Weaker links are then used to fill in missing information and to complete the propagation of these hypotheses. This polygonal search allows us to find buildings composed of multiple rectangular solids and is an improvement over techniques that are largely restricted to a single rectangular solid.

Pairwise matching using data from the first two images provides an initial graph of 3-D segments. The relationships between corners and lines in the graph are successively updated and completed with the acquisition of each new image. Each new viewpoint adds object surfaces not previously seen and provides additional observations of existing surfaces. These additional observations add information since there are usually significant changes in illumination and viewing geometry. The accumulation of multiple views also reduces potential mismatching because of accidental alignments and increases the 3-D positioning accuracy of derived object models by simultaneous solution of the collinearity equations. Surface detection can be performed on the updated graph after each step of the process or it can be deferred until the graph contains information from all of the available imagery.

3.1.1 Building generation

Once we have generated all plausible surfaces, we still need to choose the subset that can be considered as a part of a building roof. We consider two types of building hypotheses: peak roof and flat roof buildings. Once all plausible surface hypotheses are generated we perform the following steps in order to generate a set of building hypotheses:

- Radiometric pruning. Look for surfaces with nearly homogeneous image intensity distributions.
- Surface classification. Decide whether the surfaces can support a peaked roof or flat roof building hypothesis on the basis of geometric consistency.
- Height estimation. Estimate the height above ground of the building, using the vertical building segments generated by the corner matching.
- Geometric fitting. Using the set of geometric constraints provided by the building type, perform a best fit match of the building to adjust the 3-D shape.

3.2 Building generation performance evaluation

We present a representative set of results on six views (two near-nadir, four oblique) from the RADIUS modelboard imagery. Figure 5 shows the full 3-D buildings generated from the six modelboard images, using the DEM to compute height estimates, while Figure 6 shows those buildings projected to another view. We evaluated the system results using both a previously-developed 3-D voxel metric [McGlone and Shufelt, 1994a] and a new similarity metric. A broader discussion of evaluation issues and more detailed results on this and other data sets can be found in [Roux *et al.*, 1995].

3.2.1 Similarity metric calculation

The steps required to compute the similarity statistics between a building hypothesis and the set of ground truth buildings are as follows:

- **Coordinate Conversion.**

  Both the building hypotheses and ground truth buildings are transformed to the same local coordinate frame in object space.

- **Measuring the dimensions of the building hypothesis.**

  A set of unique dimension vectors are computed for each of the generated building hypotheses.

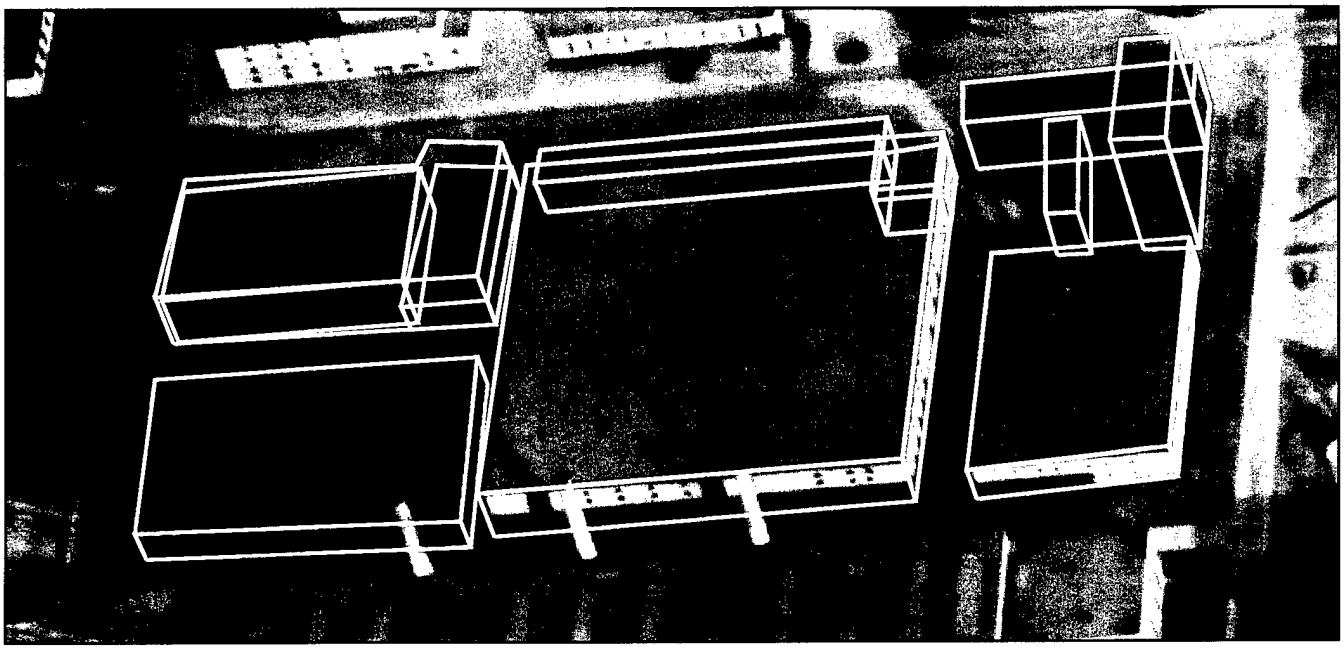


Figure 5. Buildings extracted from six modelboard images, with heights generated using a DEM.

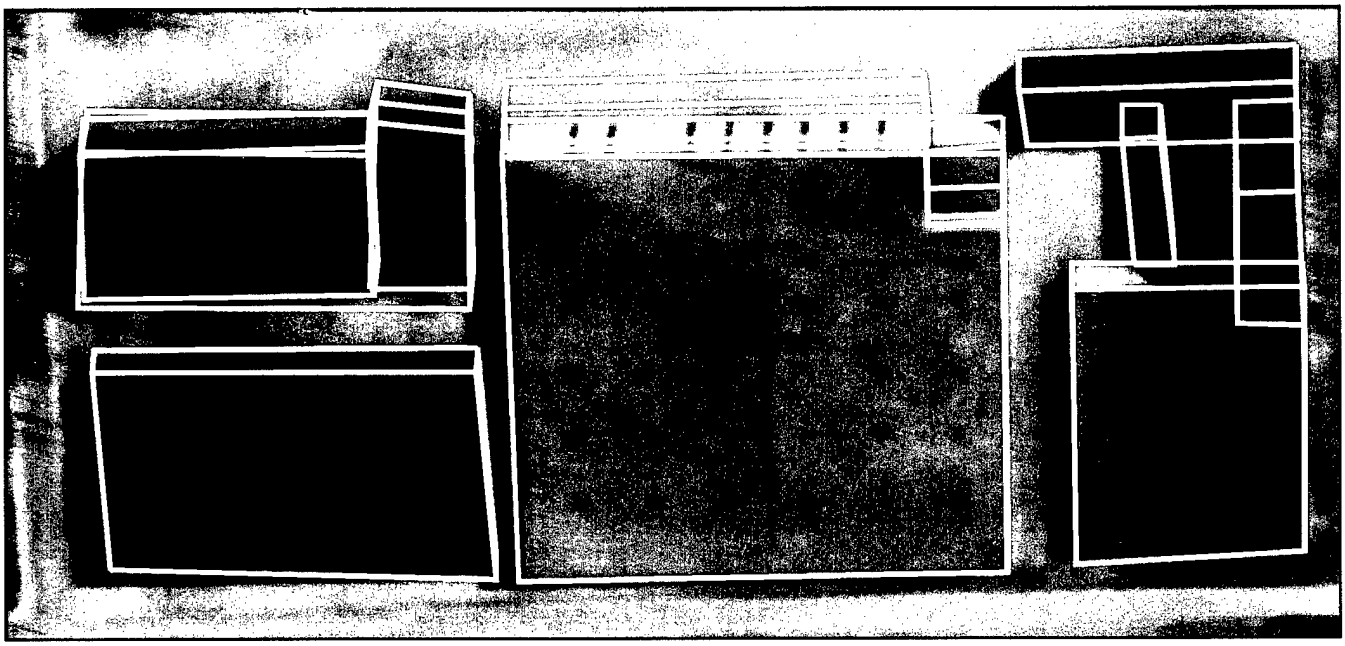


Figure 6. Extracted buildings projected into image K3, not used to calculate buildings.

- **Match building hypothesis to ground truth buildings.**

  Matched buildings must be of the same type (flat or peaked roof), have the same shape, and the hypothesis must overlap only one ground truth building by at least 50%. The sum of the length errors is accumulated, and the match with the smallest total error is selected.

- **Compare dimensions of matched buildings.**

  Once the buildings are matched, we can compute the vector differences in term of position, length, and orientation error.

3.2.2 Experimental results

The building generation results for permutation {J25,J24, J2,J4,J8,J6} are used for the evaluations described in this section. Table 1 presents the average absolute dimensional and position errors (in meters), and the orientation errors (in degrees), for six of the buildings in the scene that were detected in four, five, and six views. In this case one can observe that the physical estimates of the building shape and position remain relatively unchanged as the number of views increases. One might expect that such measurements should improve as more information is integrated. However, this conventional wisdom is not always borne out in our experiments.

Figure 7 shows the number of matched building hypotheses and the number of unmatched building hypothesis plotted against the total number of image views. This graph is typical of many of our experimental results. As the number of images is increased, more buildings can be correctly detected. However, this increase in performance is generally coupled with the generation of additional false buildings. Further, with more images, the overall accuracy of the building delineation with respect to the ground truth does not uniformly appear to improve. In some cases we do achieve improved length, width, and positional accuracy. Nonetheless, our expectation that additional images should uniformly allow for more robust matching can not be substantiated. In most cases the magnitude of the measured errors remains constant over additional views; however it does not consistently change in predictable ways. This may be explained by noting that while additional imagery provides additional information, it also can introduce more noise in all phases of hypothesis generation.

3.3 The effects of ordering permutations

In certain situations in automated man-made feature extraction we have many images available, taken over time, from which we would like to select a useful subset to actually process. Also, given a system that is capable of multiple image analysis, are there any principles of selection that can be derived by observing system performance? To evaluate the robustness and sensitivity of our system to the sequence in which images are analyzed and results combined to generate building hypotheses, a simple experiment was devised and carried out. For the six model board images used in this report, we randomly selected 50 of the 720 possible image order permutations and evaluated system performance on each of these permutations.

3.3.1 Permutation evaluation using matched building criteria

The result from each permutation was evaluated using the methods described in Section 3.2. Figure 8 shows the number of building hypotheses that were matched, building hypotheses that were unmatched, and buildings in the ground truth for which no hypothesis was generated. This table is sorted by the number of matched hypotheses and plotted against the individual permutations. Figure 8 shows that in

Table 1. Average absolute error (in meters/degrees) for 6 buildings as a function of number of views.

Nbr views	Length Error	Width Error	Height Error	Position Error	Orient. Error
4	2.92	2.25	3.17	3.83	1.03°
5	3.42	1.67	2.57	3.68	0.76°
6	3.17	2.63	2.97	3.70	1.18°

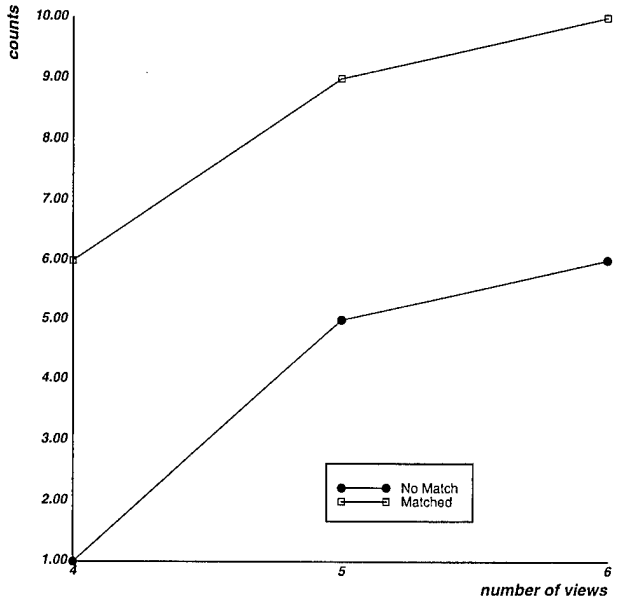


Figure 7. Matched hypotheses vs. number of views.

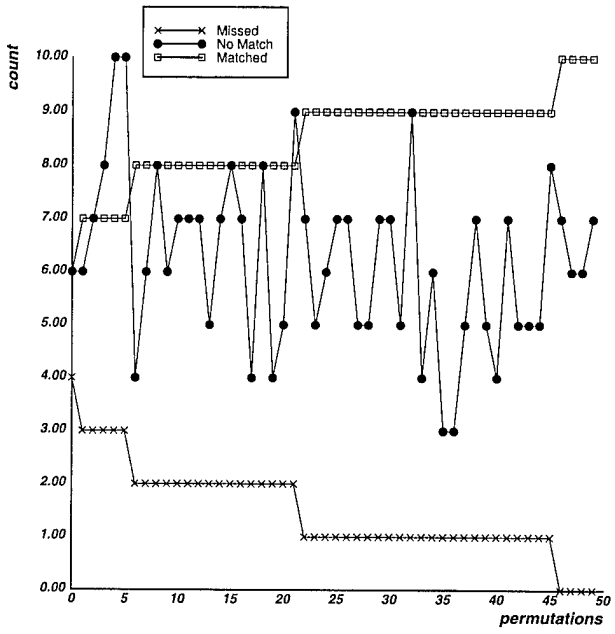


Figure 8. Matching results for 50 permutations.

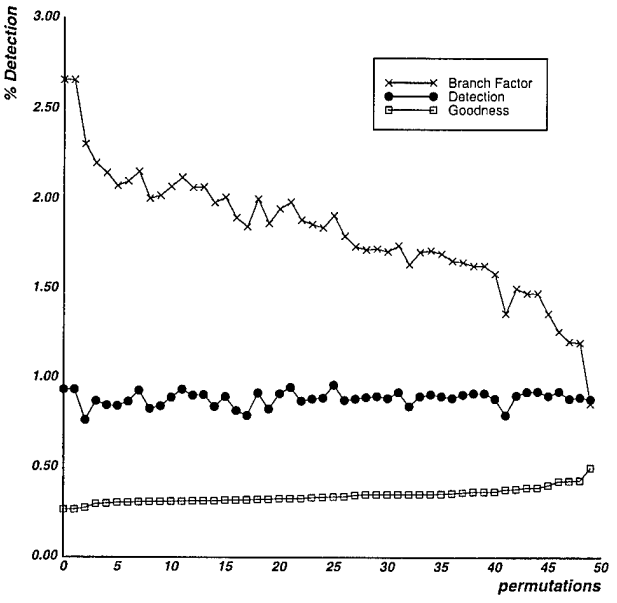


Figure 9. Voxel evaluation for 50 permutations.

28 out of 50 image permutations the system detects 9 out of 10 buildings. In four cases we are successful in detecting all 10 buildings. The worst detection performance is 6 buildings, which occurred in just one test permutation.

The first immediate observation from Figure 8 is that the number of false hypotheses is of the same order of magnitude as the number of actual buildings in the scene. This is fairly constant, regardless of the image permutation order. This generation of false hypotheses is consistent with previous work in building hypothesis generation, where independent verification techniques are required to prune incorrect hypotheses. However, it should be noted that monocular building detection and delineation systems often produce a number of false positives that is many times the number of buildings in the scene. That the number of false positives is so low can be attributed to the use of multiple images and the requirement that image cues be supported across multiple views. The overall detection performance, as well as the limited number of false positives generated, suggests that the system is fairly robust but is still susceptible to the image permutation.

3.3.2 Permutation evaluation using voxel method

Results of the voxel-based evaluation of 50 permutation runs is shown in Figure 9. Each voxel in object space is classified as true positive (TP), false positive (FP), true negative (TN), or false negative (FN) with respect to the building ground truth models. All the building hypotheses that can not be *matched* to a ground truth building are considered false positives. For building hypotheses that are matched, the voxel classification follows the metrics and methods outlined in [McGlone and Shufelt, 1994a]. The three metrics are:

- *Detection percentage* $= (100 \times TP)/(TP + TN)$.
- *Branch factor* $= FP/TP$.
- *Goodness* $= (100 \times TP)/(TP + FP + TN)$.

The permutation order in Figure 9 is sorted by the goodness measure rather than by the number of matched buildings as in Figure 8. One can see that the detection rate is nearly constant. Since all the building hypotheses that were not matched are counted as false positive voxels, the goodness measure mirrors the overall branching factor. This reflects the lack of a building verification component in the MULTIVIEW system.

Analysis of the results suggests that permutations that contain pairwise orderings of stereo pairs result in fewer false hypotheses, probably because of the stereo pairs being easier to match. This leads us to hypothesize that the most effective ordering of images is to order imagery as stereo pairs.

4 SiteCity—Semi-Automated Site Modeling

Until recently, research at the Digital Mapping Laboratory in feature extraction has addressed only automated methods. Semi-automated methods were not seen as feasible, since the low performance obtainable from previous automated methods would not reliably increase the operator's productivity. We now feel that improvements in automated algorithms, such as those used in VHBUILD [McGlone and Shufelt, 1994a; McGlone and Shufelt, 1994b], PIVOT (Section 2), and MULTIVIEW (Section 3), permit their incorporation into a highly productive and efficient semi-automated site modeling system.

The paper [Hsieh, 1996] describes SiteCity in detail, a system that combines an interactive site modeling capability with many of the algorithms developed for our work in automated building extraction, built

upon a rigorous multi-image photogrammetric foundation. Our research addresses three equally important aspects of semi-automated systems: the implementation of the automated processes to perform the intended tasks, the design of the user interface and interaction modes, and the rigorous evaluation of the system to determine if it is really an improvement over strictly manual methods.

The operational flow of SiteCity is completely flexible and under the operator's control. To illustrate the integration of automated and manual processes, we describe a typical scenario.

The operator begins by delineating the outline of a building roof in one image (Figure 10). If the building has a peaked roof, the system can automatically find the peak (Figure 11) by searching within the roof outline. The height of the building is determined by searching for vertical building edges at the roof corners (along a line toward the vertical vanishing point) and the floor lines at the bottom of the building (Figure 12). If any of these automated processes fail, the operator can adjust the height of the building manually.

The next step is to locate the building in the other images. The system projects the preliminary building model from the first image into the other images, using elevations from a DEM, then searches along epipolar lines to find the precise location of the building (Figure 13). Search bounds around the epipolar lines are determined using the image orientation covariance information, while the search range along the epipolar lines is set using expected elevations and building heights. Once the building is delineated in all images, a simultaneous solution is done incorporating the image measurements and the applicable geometric constraints [McGlone, 1995]. This produces the best estimate for building measurements and geometry, along with precision statistics for the building parameters. The final result is a true 3-D model that can be back-projected into each image (Figure 14) or displayed as a 3-D model (Figure 15).

Automated verification of the building in the other images is based on a combination of several different criteria; since we have a 3-D building model, we can use the full geometry of the building as well as scene knowledge to aid in the verification. Model edges are verified against detected image edges, excluding lines in the model that are hidden in that view. If no edge was detected at a location predicted by the model, the image gradient is re-examined for evidence of an edge. Predicted shadow edges cast by the 3-D model are used to search for shadow regions, to further verification.

Since the system is interactive, there are many different ways the task can be accomplished, depending on prior information (the operator can update or edit an existing site model), scene characteristics, or available imagery. For example, many scenes contain a number of identical or similar buildings. In this case, the operator can delineate a typical building, then instruct the system to find similar buildings by specifying a point contained in the similar building, a line that crosses a group of similar buildings, or an area that contains a group of similar buildings. The system then searches for matching buildings and displays the final results.

The application of automated methods to extremely complicated buildings is still an open issue. For this reason, SiteCity can also work in a completely manual mode. Figure 16 shows a typical SiteCity screen, working with two oblique images and one vertical image. The manually extracted building is superimposed on each image and is shown in a 3-D display.

Building a semi-automated system is only part of the task; there is no reason to use the system without a rigorous evaluation to determine whether the integration of manual and automatic methods is better than strictly manual methods. The evaluation process must address several questions:

- The performance of the automated processes, in terms of speed, precision, and reliability.
- How well the automated processes complete or augment manually delineated feature cues.
- The overall "usability" and productivity of the system.

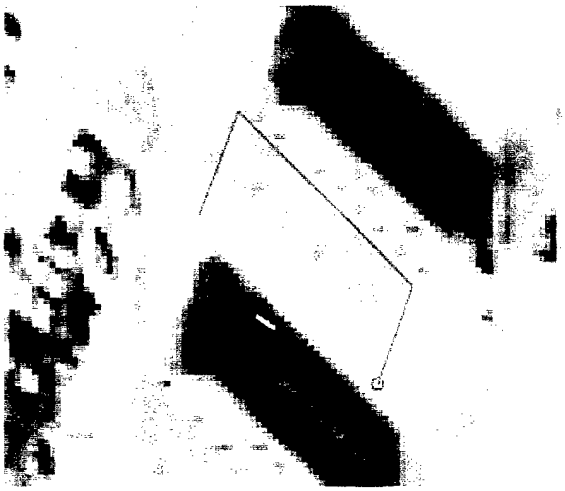


Figure 10. The user outlines the building roof.

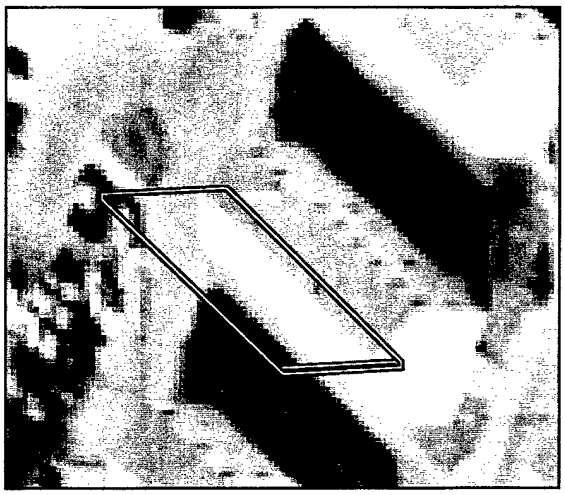


Figure 11. The system automatically finds the roof peak.

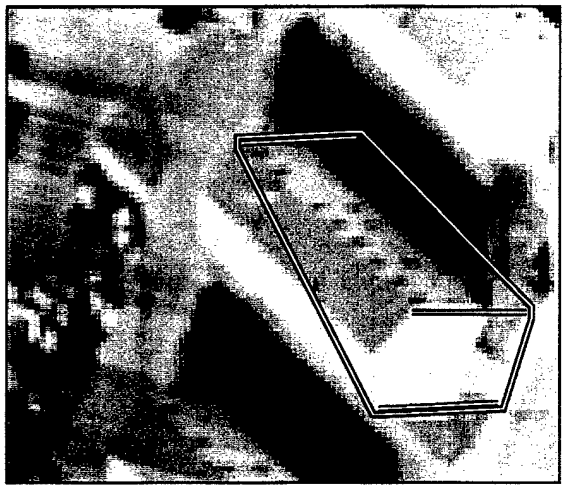


Figure 12. The system finds the floor.

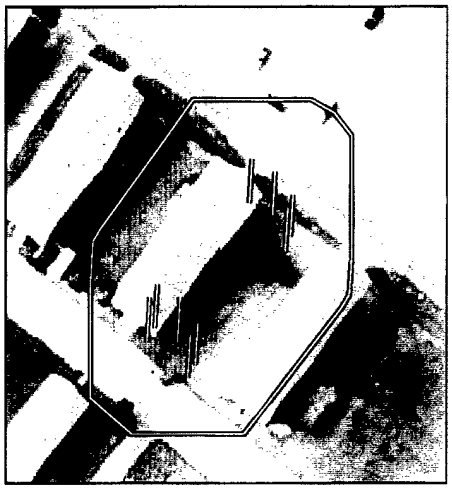


Figure 13. By searching along epipolar lines, the building is located in the other images.

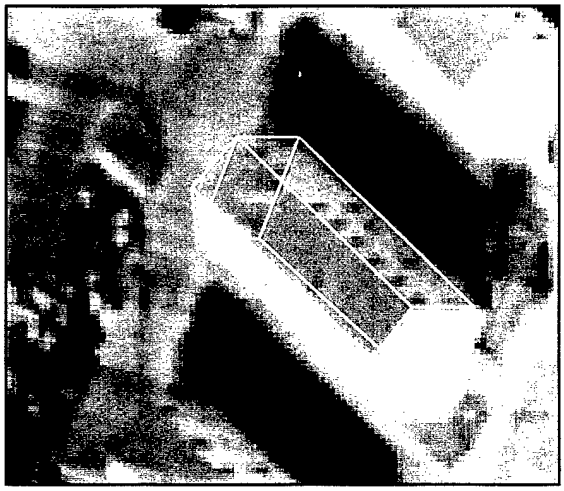


Figure 14. The final building model, projected into an image.

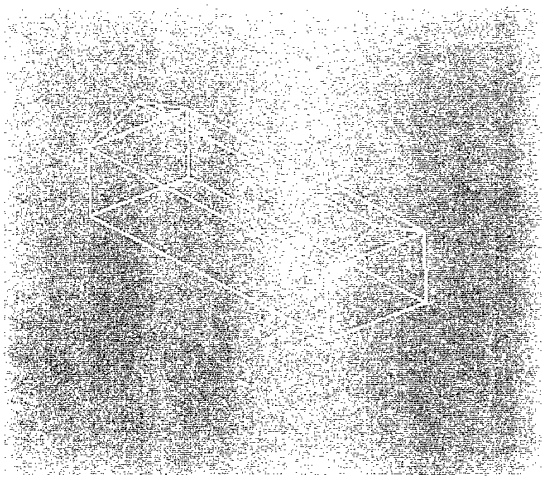


Figure 15. The final 3-D result, after simultaneous solution.

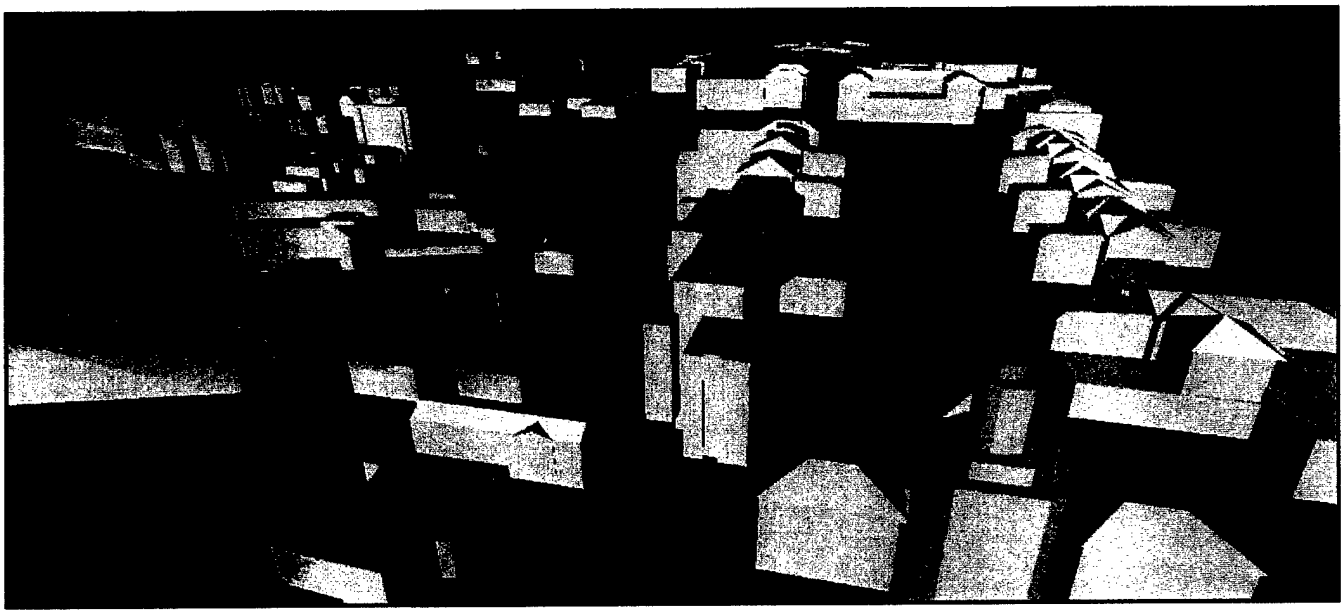


Figure 16. Model of CMU campus manually derived using SiteCity.

We have subjected SiteCity to a major evaluation effort to address these questions. Twelve test subjects measured buildings in two test scenes, manually and semi-automatically, while the operations performed and the time required were recorded. The details and results of the evaluation procedure are given in [Hsieh, 1995] and [Hsieh, 1996]; to summarize, most subjects performed the measurements more quickly and with fewer operations when using the semi-automated system. The measurement accuracy of the automated processes was shown to be as good as that for manual measurements.

SiteCity has proven to be an effective and efficient system thus far: we are currently using it to produce 3-D ground truth for our building extraction research. Current work is concentrating on the improvement of the automated processes, the ability to handle more complicated buildings, and further optimization of the user interface.

5 Object-space geometric constraints

Geometric constraints have been traditionally used to help determine sensor orientation [McGlone and Mikhail, 1982; Strunz, 1992; Mikhail, 1970; Thornton and others, 1994]; however, our main interest is the use of geometric constraints in the automated and semi-automated generation of precise site models from multiple, often oblique, images. Constraints allow the rigorous incorporation of geometric knowledge about man-made features into the solution, producing a more precise model of the site. As described in Section 4, geometric constraints are an important component of SiteCity, our semi-automated site modeling system. This additional information also enables us to more reliably test and edit assumptions about object space geometry made by automated applications, an especially valuable capability when dealing with noisy automatically extracted features.

5.1 Implementation

Our implementation of geometric constraints is part of our general photogrammetric package [McGlone, 1992; McKeown *et al.*, 1994]. A constraint can be written between any number of images, points, or other

Table 2. Comparison of building parameters and precision (meters) with and without constraints.

Param	Without Cons		With Cons	
	Value	Std. Dev.	Value	Std. Dev.
Length	219.28	0.74	219.14	0.57
Width	170.54	0.93	170.50	0.65
Height	26.63	2.01	26.95	1.25

constraints. Currently implemented constraints are of three basic types:

- Object space geometric constraints:
  - Collinearity, involving any number of points. The line orientation or position can be fixed by specifying parameter standard deviations.
  - Coplanarity, involving any number of points. The plane orientation or position can be fixed by specifying parameter standard deviations.
  - Angle between three points.
- Shadow geometry:
  - Shadow of a vertical edge (top point visible).
  - Shadow of a horizontal edge.
- Relationships between objects:
  - Same slope (lines or planes).
  - Parallel (lines or planes).
  - Same point.

Several more constraints are planned, including relative constraints on distance and azimuth, and constraints on image parameters.

5.2 Constraint applications to building modeling

Figure 17 shows one of three oblique views of a rectangular building on a modelboard, taken as part of the RADIUS program [Gee and Newman, 1993]. The eight building corner points were used in a constrained solution; the bottom points 0, 1, 2 and 3 were beneath top points 4, 5, 6, and 7, respectively. Coplanarity constraints were applied to the top and bottom planes of the building, all horizontal corners were constrained to be right angles, and the points at each corner were constrained to be vertically aligned. The building corner points were solved in a simultaneous bundle adjustment, both with and without the constraints applied. (The three images had been previously resected; their parameters were included as part of the solution, along with the covariance information from the earlier solution). The length, width and height of the building were calculated by fitting a rectangular prism to the corner points calculated with and without constraints, taking into account the covariances of the points. The derived parameters and their standard deviations are shown in Table 2. While the changes in the parameters themselves are relatively small, the improvements in precision are noticeable, particularly in building height.

5.3 Constraints and reliability

One of the most promising properties of adjustments using constraints is their much improved ability to identify erroneous measurements—their increased *reliability.* This is especially important for applications based upon automated feature extraction and matching, where bad matches can occur frequently and may have extreme effects on the results of the algorithm.

Equally important for automated algorithms is the ability to detect a bad constraint—to recognize and edit an erroneous geometric hypothesis. Standard reliability theory [Förstner, 1987] deals with the detection of blunders in measurements; by using the unified approach to least squares [Mikhail, 1980], we can extend standard reliability theory to constraint equations. In the unified approach, constraints are treated as observation equations with a given *a priori* weight, and containing a fictitious observation. By calculating the relability statistics for these fictitious observations, we can determine if the constraint "observation" contains a blunder—in other words, if the equation is invalid. The details of the derivation of reliability statistics for constraint equations are given in [McGlone, 1995].

In an automated application we envision the system applying the constrained solution to a hypothesized geometric situation. By examining the residuals and statistics from the solution the system will be able to verify the hypothesis and to identify faulty elements or matches. Of course, the same considerations in testing residuals apply as in the classical case [Förstner, 1994]. The ability to isolate bad observations is completely dependent upon the redundancy and the geometry of the adjustment.

5.4 Examples

To illustrate the properties of adjustments with geometric constraints, the same data set used in the previous section was re-run with a bad image measurement (the column measurement of point 6 on image J24 was increased by 10 pixels). The tau statistic [Pope, 1975] was used to determine residual editing criteria, since we are interested in examining the largest residual(s) in a rigorous manner. The plots in Figures 18, 19, and 20 show the standardized residuals for each point on each image, reprojected into a vertical image of the scene and combined for comparison purposes.

Comparing Figure 18, without any geometric constraints, and Figure 19, with constraints, shows that the bad measurement has been isolated, in that its standardized residual is much larger than the rest of the residuals. Additionally, it is the only standardized residual greater than the tau rejection criterion.

These results are hardly unexpected—adding redundancy and geometric strength to a solution will always improve its resistance to bad data. The question is, how resistant is the solution to bad constraints? Figure 20 shows the residuals from an adjustment where point 6 was replaced by a point approximately 43 meters away (about 80 pixels on image J5), again measured on all three images. This represents a common case in building matching across images, where a point is mistakenly identified as a building corner. The three right angle constraints involving point 6 were therefore invalid, since the real angles were not actually right angles.

While the image residuals of all the points involved in constraints with the bad point were bad, the erroneous point (number 6) has the largest standardized residuals. Examining the residual plot (Figure 20) shows the correlation between the residuals because of the bad geometric information, especially at the erroneous point where all the residual vectors point in roughly the same direction. The standardized residuals on the constraint observations are also bad (Table 3), although their values are harder to interpret. The largest residuals are on the bottom plane constraint, which does not directly involve point 6. Instead, the distortion because of the bad point 6 is spread through the geometric figures as the solution tries to

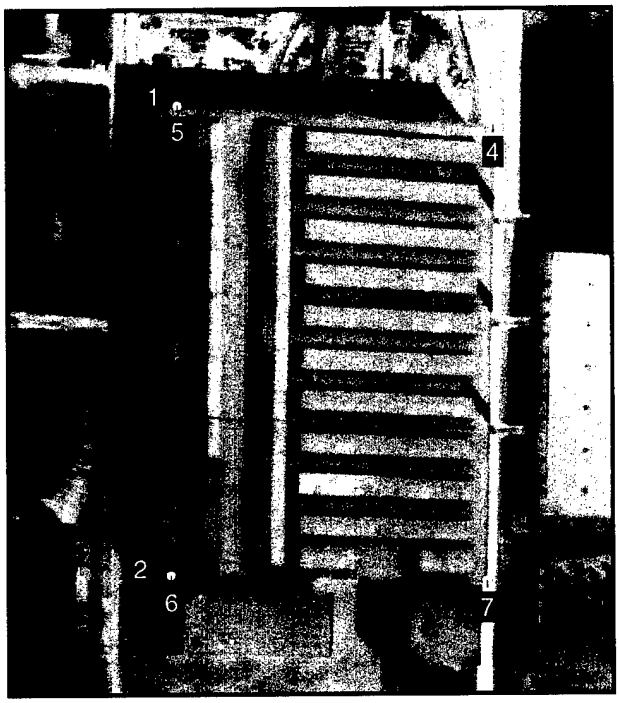


Figure 17. Image J5.

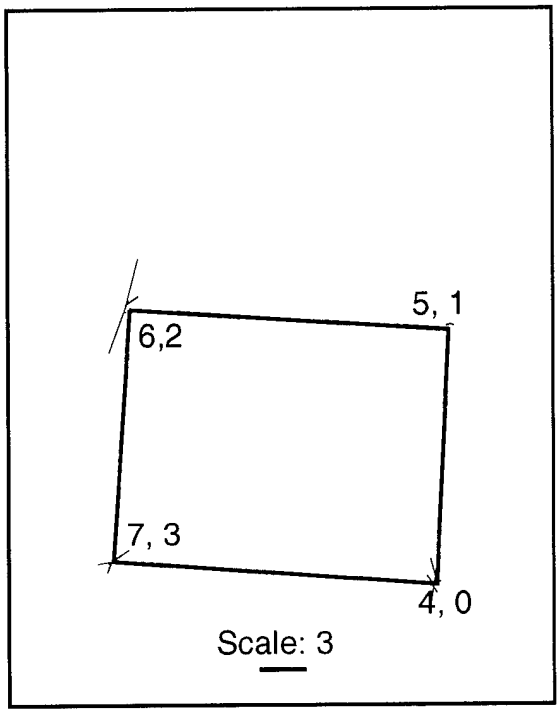


Figure 18. Standardized residuals from solution with no geometric constraints.

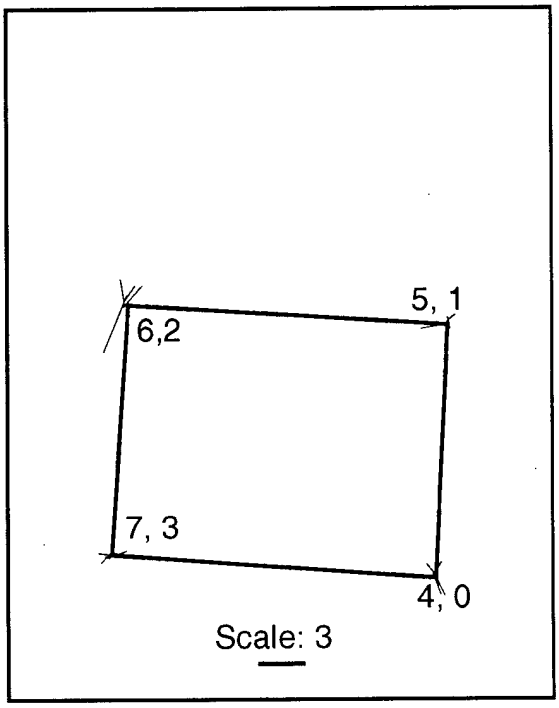


Figure 19. Standardized residuals from solution with geometric constraints.

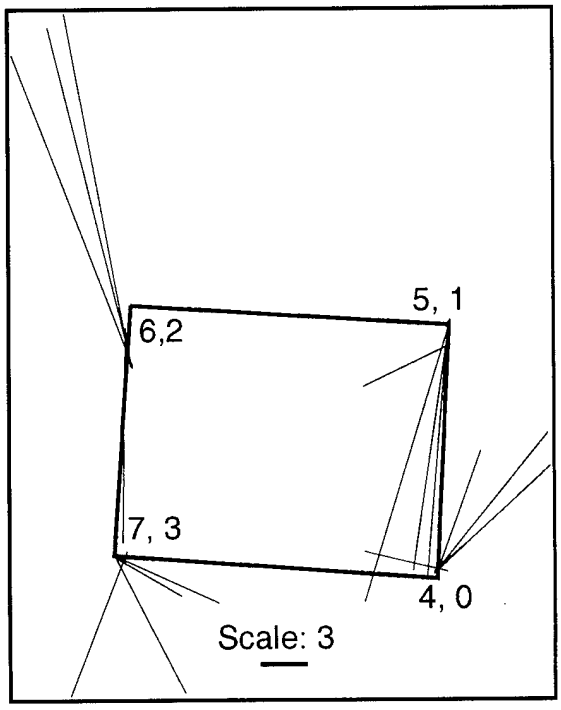


Figure 20. Standardized residuals from solution with bad constraint.

Table 3. Constraint equation standardized residuals for case with invalid constraint.

Constraint	Points	Std. Residual
top plane	4	4.7
	5	-3.2
	6	-1.9
	7	1.6
bottom plane	0	-6.1
	1	0.4
	2	17.7
	3	-11.2
angle 1	at 5, from 4 to 6	-3.3
angle 2	at 6, from 5 to 7	3.7
angle 3	at 7, from 6 to 4	1.5
angle 4	at 4, from 7 to 5	-1.9
angle 5	at 1, from 0 to 2	-3.3
angle 6	at 2, from 1 to 3	3.7
angle 7	at 3, from 2 to 0	1.6
angle 8	at 0, from 3 to 1	-1.9

satisfy the angle constraint by moving the points out of the horizontal planes. The largest residual is for the equation constraining point 2 (constrained to be directly below point 6) to the bottom plane.

Further research is being done in developing this approach to applying and editing geometric constraints.

6 Road Network Extraction

The automatic extraction of roads from aerial imagery has been an active research subject in computer vision for over a decade. Approaches have varied from multispectral analysis to linear feature detection to structural analysis. Most often the distinction between detection and delineation has not been made explicit—that is, finding roads and accurately describing their properties are treated as the same problem. In our research we have explicitly divided the task of automated road network extraction into three distinct phases: road finding, road tracking, and network construction.

Our previous work in road finding used an edge-based algorithm with geometric smoothing to generate reliable starting points from which a tracker can initialize its road surface and boundary models [Zlotnick and Carnine, 1993]. Our road tracking research demonstrated the use of multiple cooperative methods, combining low-level correlation and edge-based methods with a high-level analysis component [McKeown and Denlinger, 1988]. In this section, we show some recent results from our automated road extraction system and present a new system for interactive road extraction.

6.1 Automated Road Extraction

We have continued to build on our previous research in automatic road extraction. The system as described in [McKeown and Denlinger, 1988] and [Zlotnick and Carnine, 1993] has undergone several changes. Several new edge-based methods for road finding are being used, and we have augmented the road finding and

tracking systems to use parameters defined in object-space. In addition to facilitating our move toward an object-space road tracking system, this latter change enables the system to perform road finding in oblique images and allows the same set of physical parameters to be used for most images.

Some recent results from this system are shown in Figure 21. This is a USGS digital orthoquad image over Palatine, IF (a suburb of Chicago). The image is 8 km by 7.4 km (8008 $\times$ 7380 pixels) and has a ground sample distance of 1 meter/pixel. Our road finding system generates 2427 road seeds, and the tracking system generates 3177 road segments. Figure 21 shows the final set of road segments overlaid on the image in white.

One can see that the system does a good job of finding the primary road areas. Looking at this figure gives one a good impression for where the major thoroughfares are and where they intersect. It is equally clear that the automated system has problems in some of the suburban areas where vegetation is prevalent and obscures or fragments the road edges. For example, a number of the neighborhood roads (along the left and right edges of the image) are missed completely. On the other hand, most of the striae found among the field areas and in the vegetation do not generate false-positives because of lack of edge support for a road hypothesis. Further work on this portion of the system will attempt to augment the current edge-based road finder with a new surface-based finder.

6.2 Interactive Road Extraction

To complement our work in automated road extraction, we have begun the development of an interactive road extraction system. The system, called Idl_Woof, is built upon the automated road extraction systems and our X11-based image display library (IDL). Figure 22 shows a snap-shot of a session with Idl_Woof. In this example, we have four windows (plus an edit control window) displaying automatically extracted road seeds (window 4), a portion of output from both the edge and surface-based road trackers (windows 2 and 3, respectively), and a reduced view of a subset of the extracted roads (window 1).

The interface allows the user to direct execution of the automated road finding and tracking processes, as well as permitting manual delineation and editing of road segments. Two different modes can be used when extracting a road using the automated extraction system. First, the starting parameters for the automated tracking system (position, direction, and estimated width) can be graphically selected by placing a marker on the portion of the road to be tracked. Alternatively, the automated road finder can be applied to a selected area of the image, and the generated road seeds can be selected for use by the tracker. The tracker does make mistakes, so the user has the option to reject, accept, or edit the automatically extracted segment. The interface places no constraints on the number of windows or what data are displayed in those windows.

We have found Idl_Woof useful in producing road ground truth to be used in the performance analysis of the automatic road extraction system. It also has proven useful when debugging modifications to the road finding and tracking systems.

6.3 Toward Road Network Extraction

In the near future, we will extend this work in two main research areas. First, we will focus on improvements to our road network generation system. We want to apply knowledge of road network topology to the road tracking results to create a final road network. Such a network would be constructed in object space to aid the use of multiple data sources and 3-D model information (i.e., other feature extraction results, tilt-grade information).

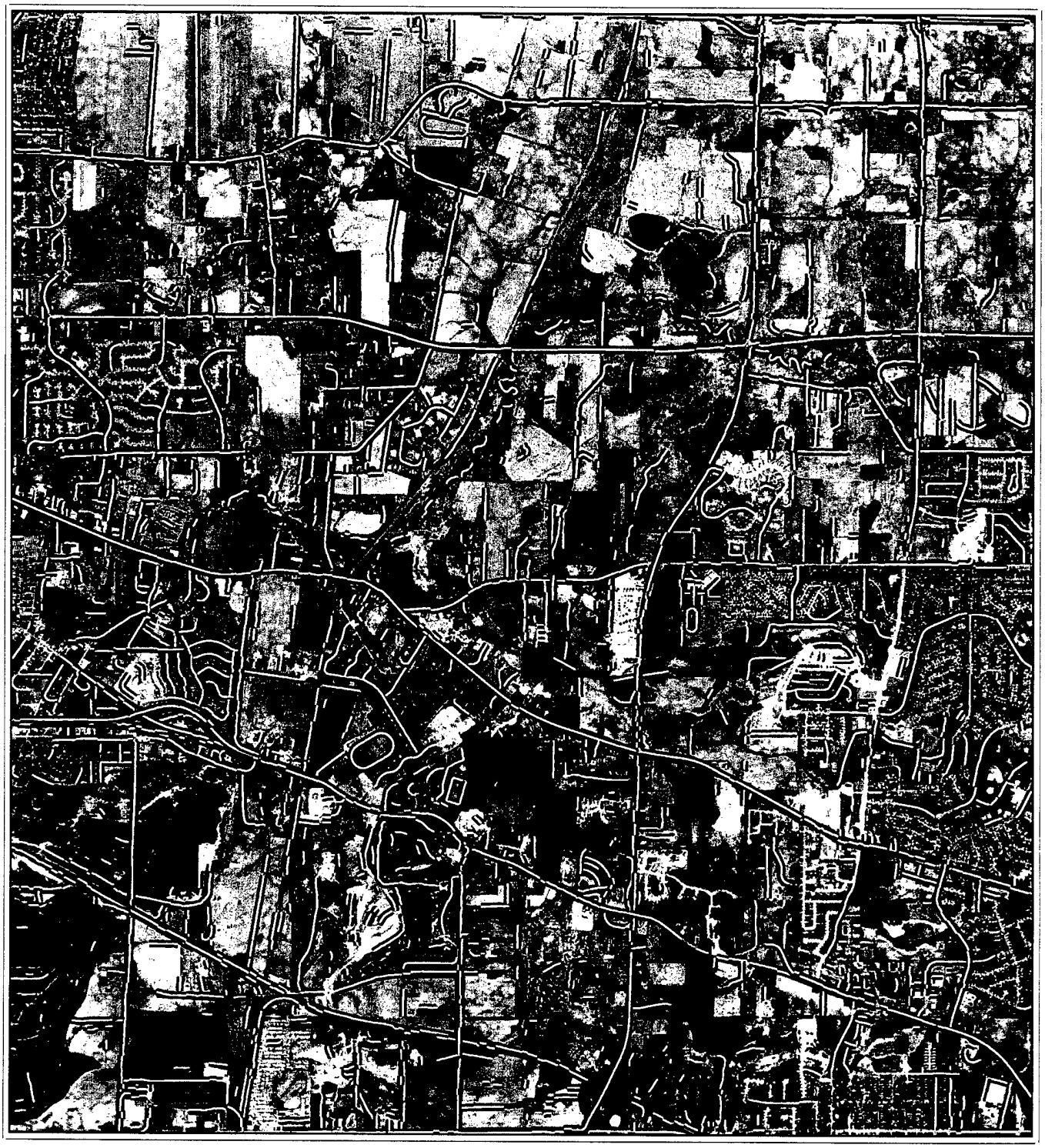


Figure 21. Automated road extraction results on USGS digital orthoquad image.

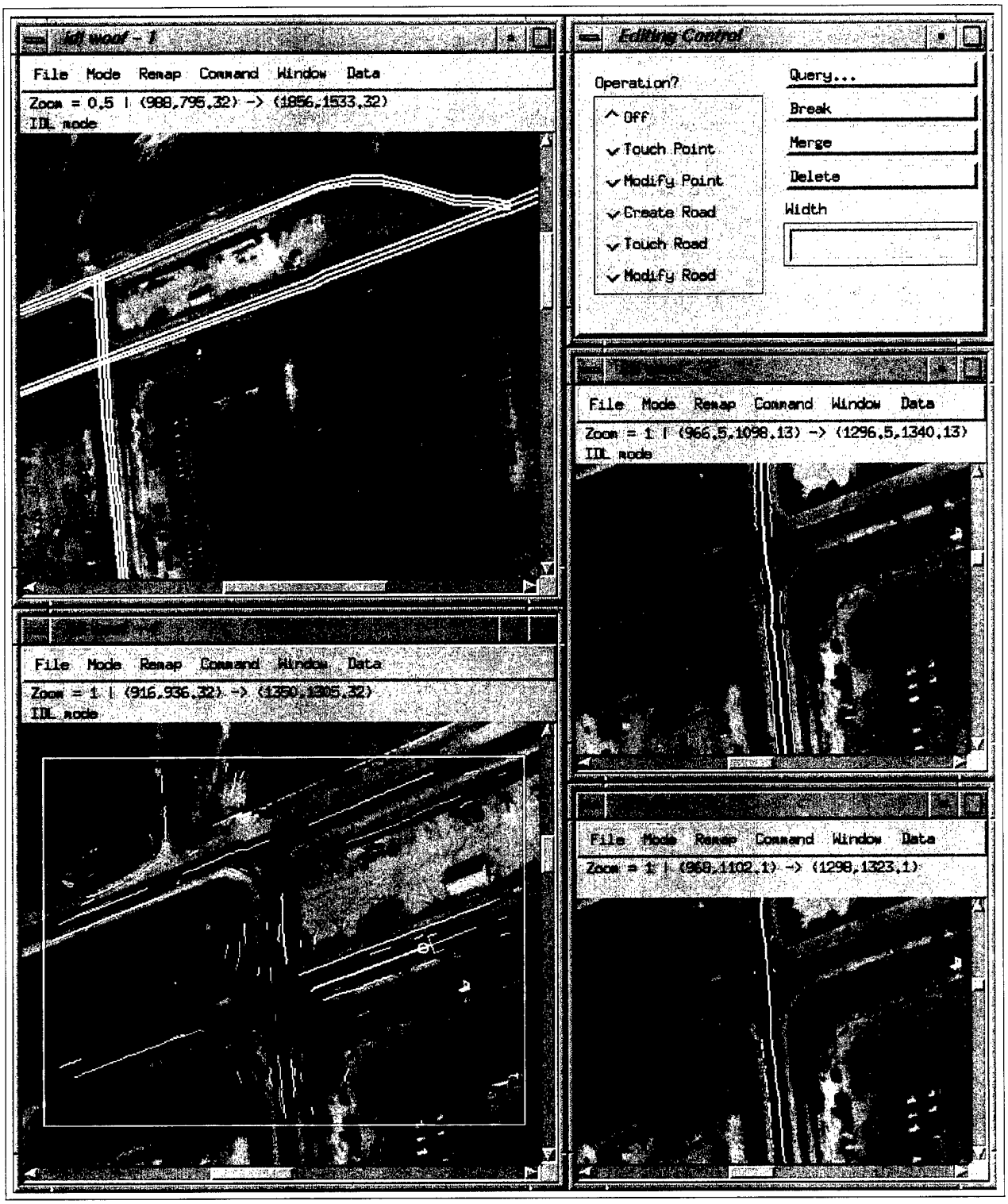


Figure 22. An example interaction with Idl_Woof showing automatically generated road seeds, output from both the surface and edge trackers, and a portion of the final output.

As a second research area, we will continue the development of Idl_Woof. In order to simplify the user interaction, we are looking at the addition of methods to automatically adjust manually-delineated roads, based on portions of the automated road finding system already in use. We also plan to include semi-automatic aids for road network construction. This will require methods for combining multiple tracked roads and an interface for annotating road features. A richer set of road feature primitives also must be added to our road model representation.

7 Stereo

In our older stereo research, the work has been directed to the stereo matching of near-nadir binocular imagery. In the recent past, we have extended that to binocular oblique imagery by adaptively adjusting the "vergence" of the stereo pair so that the actual search range was kept small [Cochran, 1994a; Cochran, 1994b; Cochran, 1995]. This maintained quick stereo matching and reduced the chance of aliasing while allowing both oblique imagery and very rugged terrain. Our current work has been to extend the stereo matching to more than two images. The first approach, MULTIVIEW used a high-level feature-based matcher and is described in Section 3. Here we describe the second system, a multiple image area-based matcher, which we call "S3."

In addition, we are enhancing our stereo capability with the import of the TEC Digital Photogrammetry Compilation Package (DPCP) [Norvelle, 1981; Norvelle, 1992] and we are in the process of retooling the man-machine interface with this program.

Finally, we have continued experiments with the data fusion of monocular and stereo data [McKeown *et al.*, 1994] with the fusion now being done fully in object-space.

7.1 S3: Multiple Image Area-Based Stereo

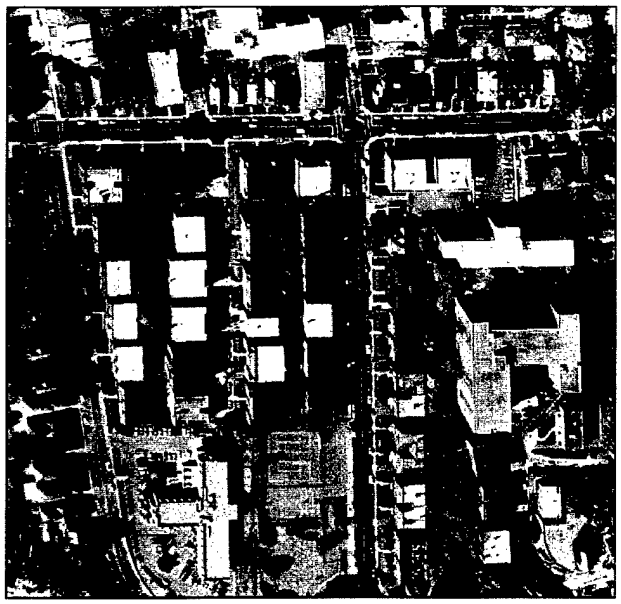
We are continuing work on the area-based stereo processing of multiple images[McKeown *et al.*, 1994]. Results are currently poor because of a combination of the need for fine local registration of the epipolars between images and the desire to avoid a combinatorial expansion as larger numbers of images are considered.

The current version calculates a multiple-image correlation score of those images that do not have a high auto-correlation. These values are computed for a correlation window existing in object-space which are projected into each of the target images, thus automatically applying image warping and sub-pixel/multiple-pixel scaling. While this works well, the process is both slow and, if the search-space is extended perpendicular to the epipolar, then the matching time increases combinatorially with additional images. We are currently working on a preprocessing phase to minimize the local mis-registration that currently exists in order to make the overall matching time increase linearly with additional images.

7.2 Import of the TEC Stereo Package

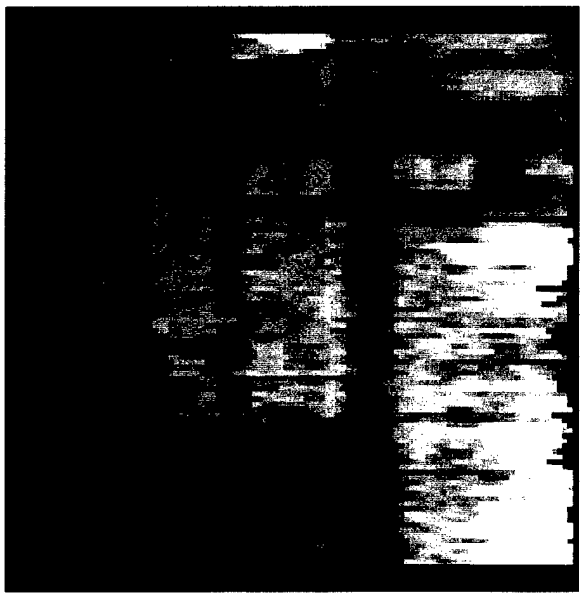
We have been working with TEC to do a technology transfer of TEC's Digital Photogrammetry Compilation Package (DPCP) [Norvelle, 1981; Norvelle, 1992]. This man-in-the-loop area-based stereo package has been used for several years at TEC for the development of Digital Elevation Models (DEMs).

We are interested in applying this stereo technique to the evaluation of aerial imagery, in particular, to cartographic feature extraction. Figures 23 and 24 show two near-nadir stereo pairs: The first, a suburb of Pittsburgh, the second, a barracked area at Fort Hood. To each image pair our older S2 stereo process (Figures 23(c) and 24(c)) were run without refinement [McKeown and Perlant, 1992], while the TEC DPCP



(a) Left image (Pittsburgh).

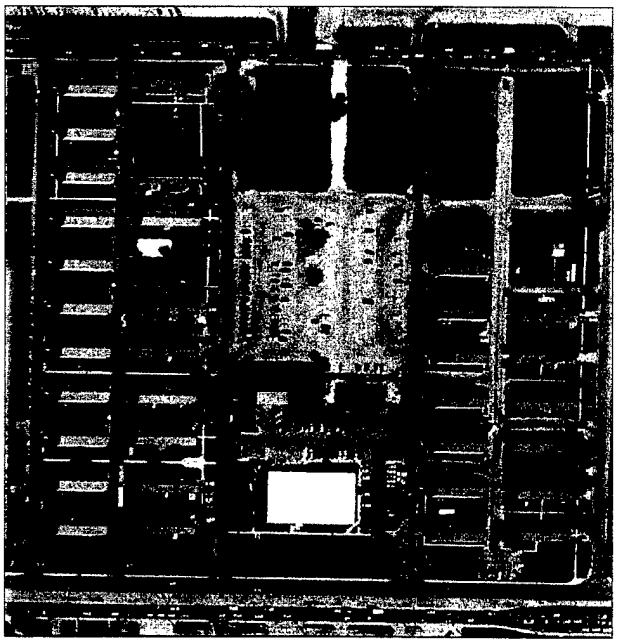
(b) Right image (Pittsburgh).



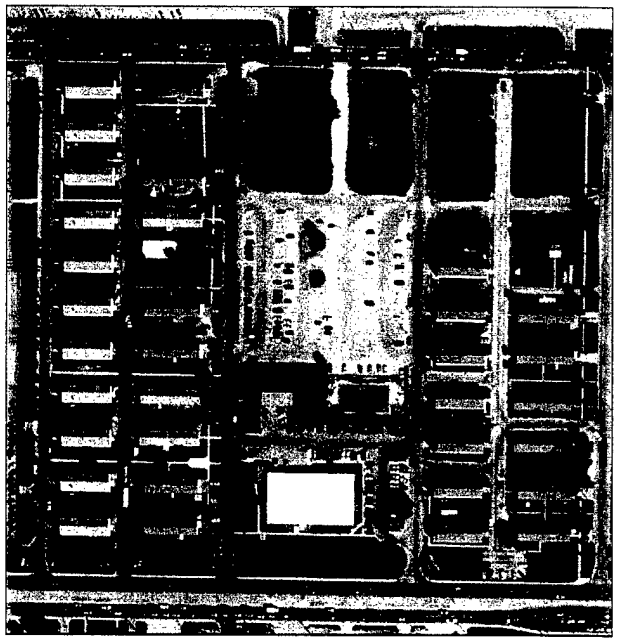
(c) S2 stereo elevation results (unrefined).

(d) DPCP stereo elevation results (without IOR).

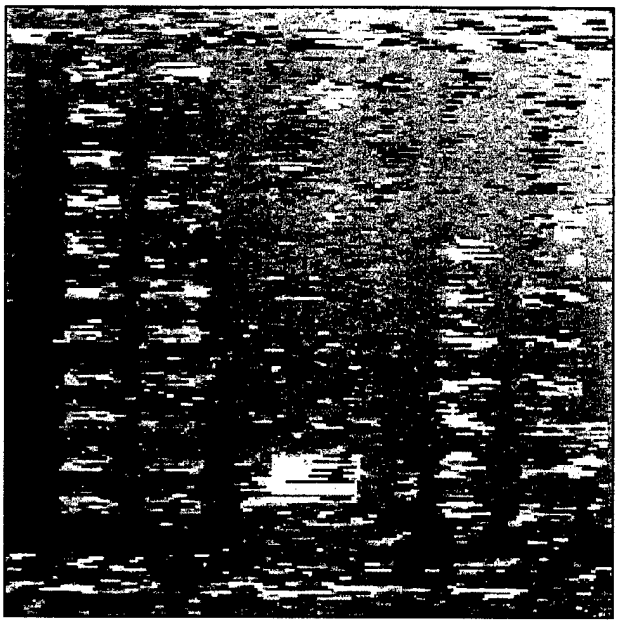
Figure 23. Raw stereo results from the S2 and DPCP processes on the Pittsburgh imagery.



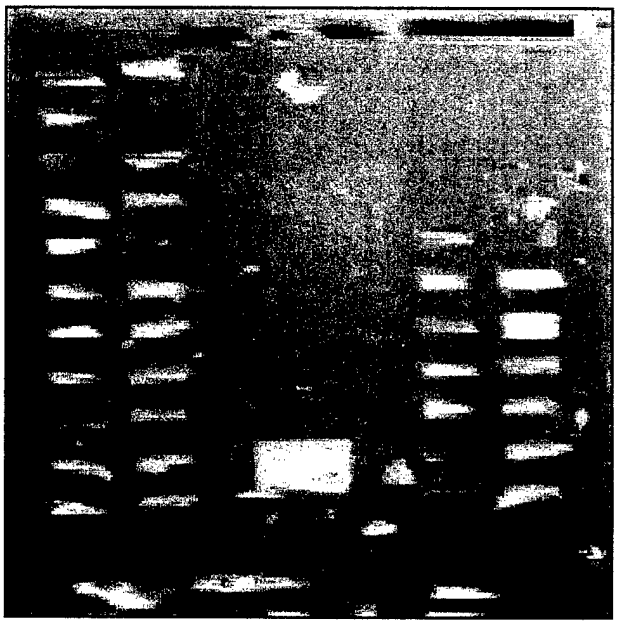
(a) Left image (Ft. Hood).



(b) Right image (Ft. Hood).



(c) S2 stereo elevation results (unrefined).



(d) DPCP stereo elevation results (without IOR).

Figure 24. Raw stereo results from the S2 and DPCP processes on the Fort Hood imagery.

(Figures 23(d) and 24(d)) examples were run without the Iterative Orthophoto Refinement (IOR) phase [Norvelle, 1992]. These results compare the raw initial processing by the stereo algorithms. Although we do not yet have a full quantitative analysis, spot checks reveal two things: First, both methods generate approximately the same elevation results when away from step edges, and second, the stronger continuity constraint built into the DPCP algorithm tends to more greatly smooth the step edges, while at the same time removes some of the noise inherent in the S2 process. However, the S2 process captures the edges and weaker edges (in terms of edge strength). Both of these processes adaptive adjust their search range and therefore work with oblique imagery.

We are in the process of further integrating the TEC DPCP into our set of tools by providing a new, more user friendly, front end. This will serve to minimize the man-in-the-loop portions of the algorithm and provide hooks for their automation. In addition, further comparison of both algorithms after S2 refinement and the DPCP IOR process will be studied and, as discussed below, these results will be integrated with other modes of information.

7.3 Data Fusion

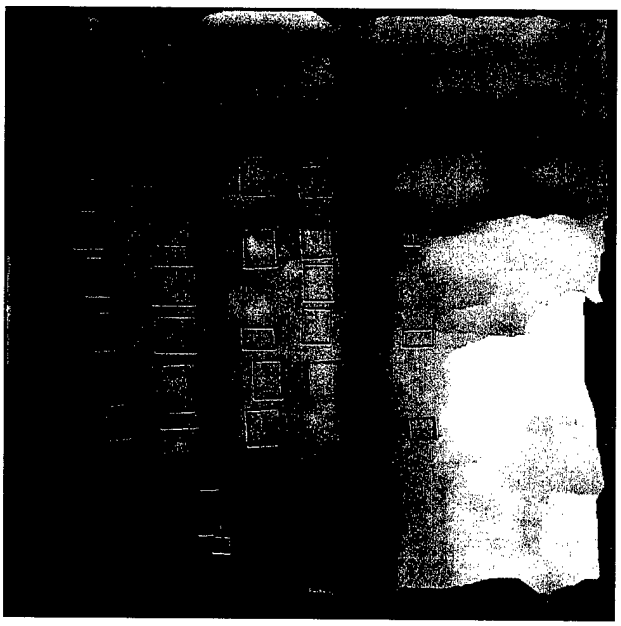
Efforts to integrate monocular and stereo data have continued (as shown below in Figure 25) for the two test scenes. Here the background is the ortho-elevation image of the DPCP elevation estimates (with brighter areas indicating higher elevations). Superimposed over this are the monocular hypotheses generated by the VHBUILD system [Shufelt and McKeown, 1993; McGlone and Shufelt, 1994b]. The hypotheses marked in red have been rejected while those marked in green have verified by checking the general elevation within the area that composes the building hypothesis against the nearby area outside that region.

We can presently use the building hypotheses along with a fusion of the stereo and the VHBUILD height estimate to generate building models in object space (see Figure 26). This works well since VHBUILD relies on vertical edges to estimate height—which works best on oblique imagery—and the stereo processes work best on near-nadir imagery. Thus, these two approaches are very complimentary. Future data fusion will include material classification generated from multispectral data where it is available (see Section 8).

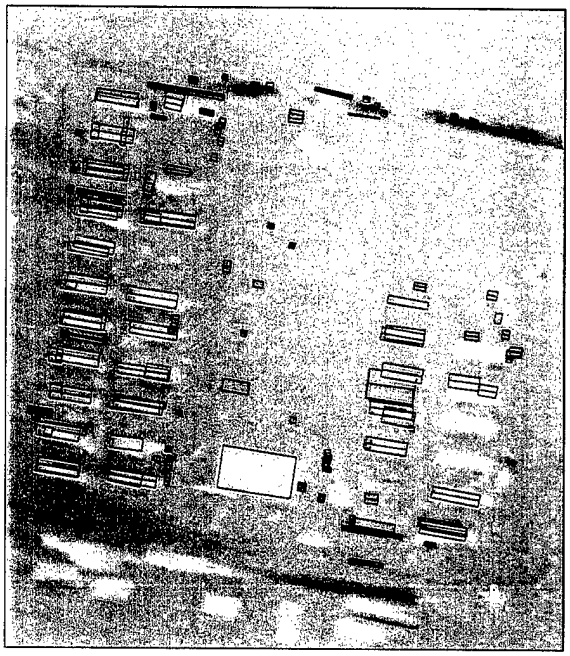
8 Hyperspectral Data Acquisition over Fort Hood, Texas

Our previous work has shown the feasibility of merging surface material information derived from moderate resolution multispectral imagery with estimates of height based upon stereo matching in high resolution panchromatic imagery [Ford and McKeown, 1992a; Ford and McKeown, 1992b]. The goal is to use surface material information, normally highly correlated with object location in complex urban scenes, as a source of information for small scale mapping of man-made structures such as buildings and roads, as well as natural features, such as soil, vegetation, and water. The fusion of height estimates with surface material estimates provides a unique synthetic 3-D dataset that is not directly available in any airborne imaging sensor. With the availability of high resolution multispectral/hyperspectral imagery, comparable in spatial resolution to aerial mapping imagery, opportunities exist to exploit the inherent spectral information of multispectral/hyperspectral imagery to aid urban scene analysis for cartographic feature extraction and simulation database population.

A window of opportunity occurred during October 1995, to collect high spatial resolution hyperspectral imagery with the Naval Research Laboratory's (NRL) HYDICE sensor system over Fort Hood, TX. Fort Hood has been a focal point for our research and experimentation in automated cartographic feature extraction (i.e. buildings and road), stereo analysis and spatial database construction. We have a variety of imagery and digital cartographic datasets presently inhouse covering portions of Fort Hood, including:



(a) Verified (green) and unverified (red) monocular building hypotheses.



(b) Verified (green) and unverified (red) monocular building hypotheses.

Figure 25. Fusion of monocular building hypotheses with the DPCP stereo results.

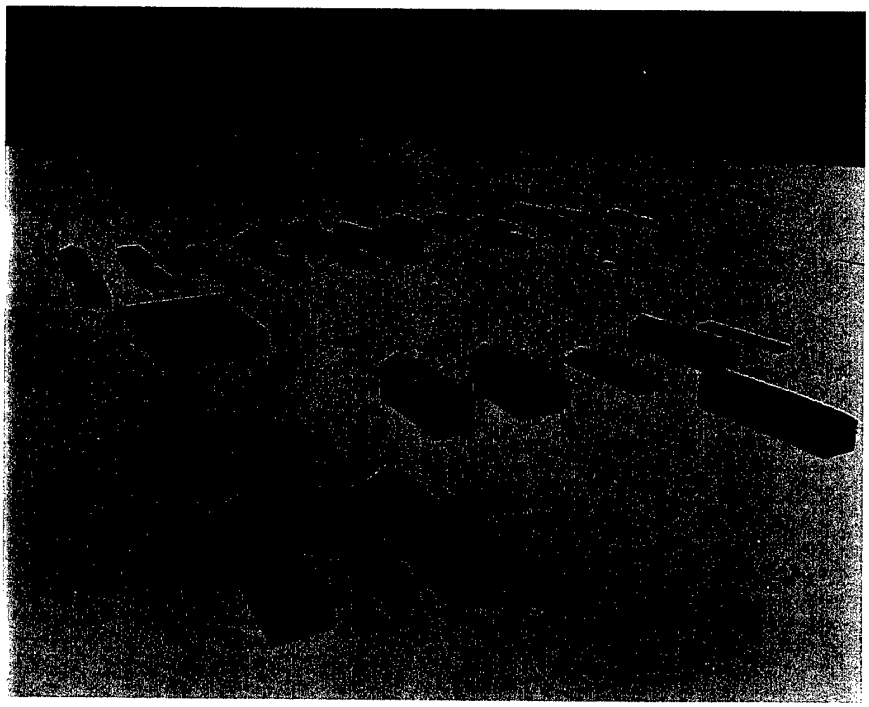


Figure 26. Ground truth model of the verified building hypotheses fused with the DPCP elevation estimates.

- 3 National High Altitude Program (NHAP) images (1.2 meter GSD).
- 9 RADIUS nadir images (0.5 meter GSD).
- 21 RADIUS oblique images (0.5 meter GSD).
- 1 SPOT multispectral (XS) image (20 meter GSD).
- DMA Digital Terrain Elevation Database (DTED) (Level 1 and 2).
- USGS Digital Elevation Model (DEM) (1:250,000 scale).
- DMA Interim Terrain Database (ITD) (1:50,000 scale).
- USGS Land Use and Land Cover (LULC) (1:250,000 scale).
- USGS Digital Line Graphs (DLG) (1:100,000 scale).

Adding a high resolution hyperspectral dataset to this collection will allow us to pursue research in surface material/land cover classification at a spatial resolution comparable to the NHAP and RADIUS imagery for high resolution mapping applications. Since each separate data set is registered to a common geodetic framework, we will be able to merge individual datasets or the outputs of our cartographic feature extraction systems to produce more complex cartographic products.

Readily available digital cartographic data sources, such as USGS LULC and DMA ITD, do not support this high spatial requirement. Features such as roads, parking lots, buildings, tree canopies and grass areas are quite evident in Figure 27, a portion of a mapping image collected over the motor pool and barrack areas of Fort Hood. The corresponding area represented in the DMA ITD and USGS LULC cartographic datasets aggregates these features together as shown in Figures 28 and 29.

We will perform experiments with HYDICE imagery that attempt to replicate our previous results using Daedalus scanner data in the Washington D.C. area [Ford and McKeown, 1992a; Ford and McKeown, 1993; Ford *et al.*, 1993], using Fort Hood as the test site. We expect significantly improved stereo results given higher resolution mapping photography, improved matching algorithms. Likewise we expect a major improvement in sensor registration and surface material identification using the higher spatial and spectral resolution available with the HYDICE scanner. This synthetic 3-D dataset will be used to augment and intensify a Fort Hood cartographic database initially constructed using DMA ITD and USGS LULC.

8.1 HYDICE Sensor System

The Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor system is mounted on a CV-580 aircraft; depending on aircraft altitude above ground level, the ground sample distance (GSD) varies from 1 to 4 meters. The HYDICE sensor is 320 pixels wide, giving a ground swath of 320 meters up to approximately a kilometer. The spectral range of the HYDICE sensor extends from the visible to the short wave infrared (400 to 2500 nanometers) region, divided into 210 channels with nominal 10 nanometer bandwidths. Figure 30 illustrates the HYDICE sensor spectral bandpasses with respect to a schematic version of the average atmospheric transmission curve from the earth's surface to the top of the atmosphere. The bandwidths of HYDICE vary from 7.6 to 14.9 nanometers, depending on channel location in the electromagnetic spectrum. Additionally, the spectral bandpasses of three multispectral imaging systems (Daedalus Airborne Thematic Mapper (ATM), Landsat Thematic Mapper (TM) and SPOT High Resolution Visible (HRV) Imaging Instrument) are shown to demonstrate the high spectral resolution of the HYDICE sensor system.

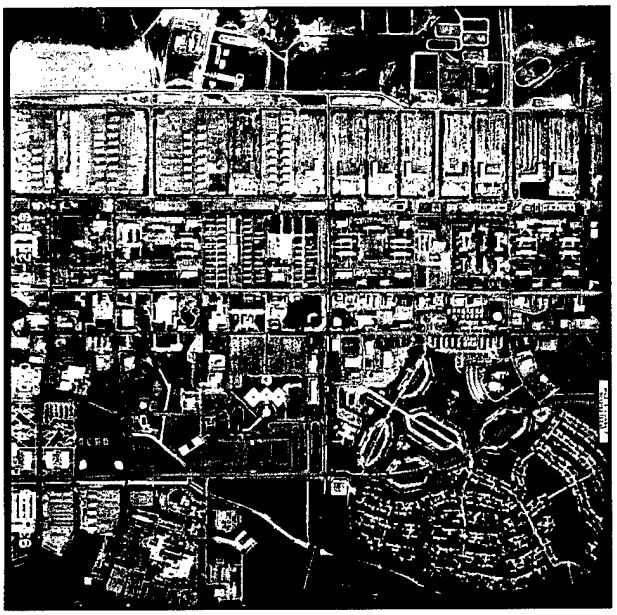


Figure 27. Mapping image FHN713.

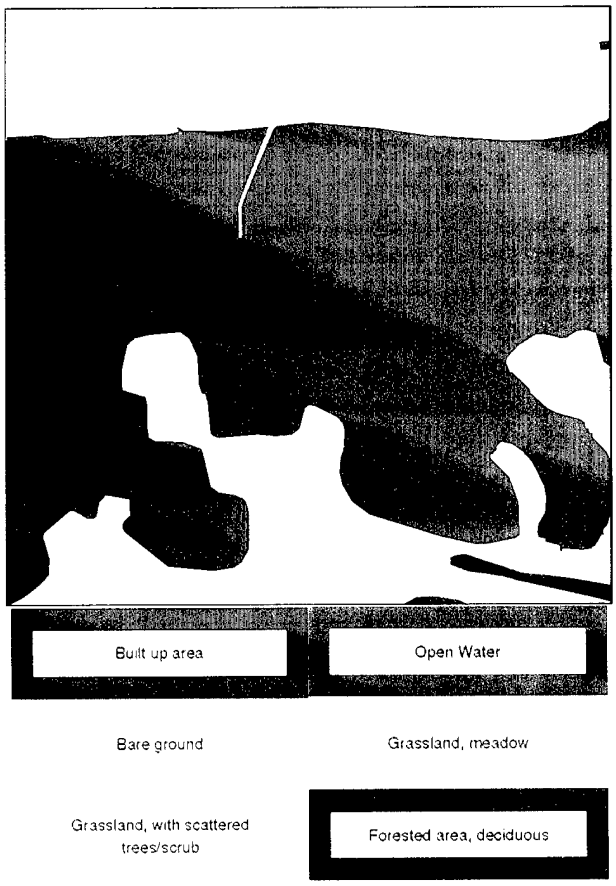


Figure 28. DMA ITD vegetation coverage with FHN713.

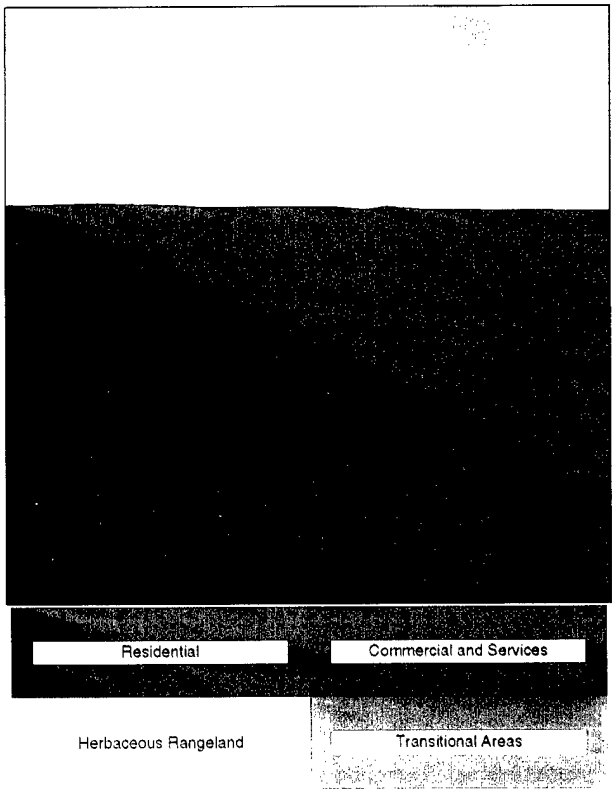


Figure 29. USGS LULC coverage with FHN713.

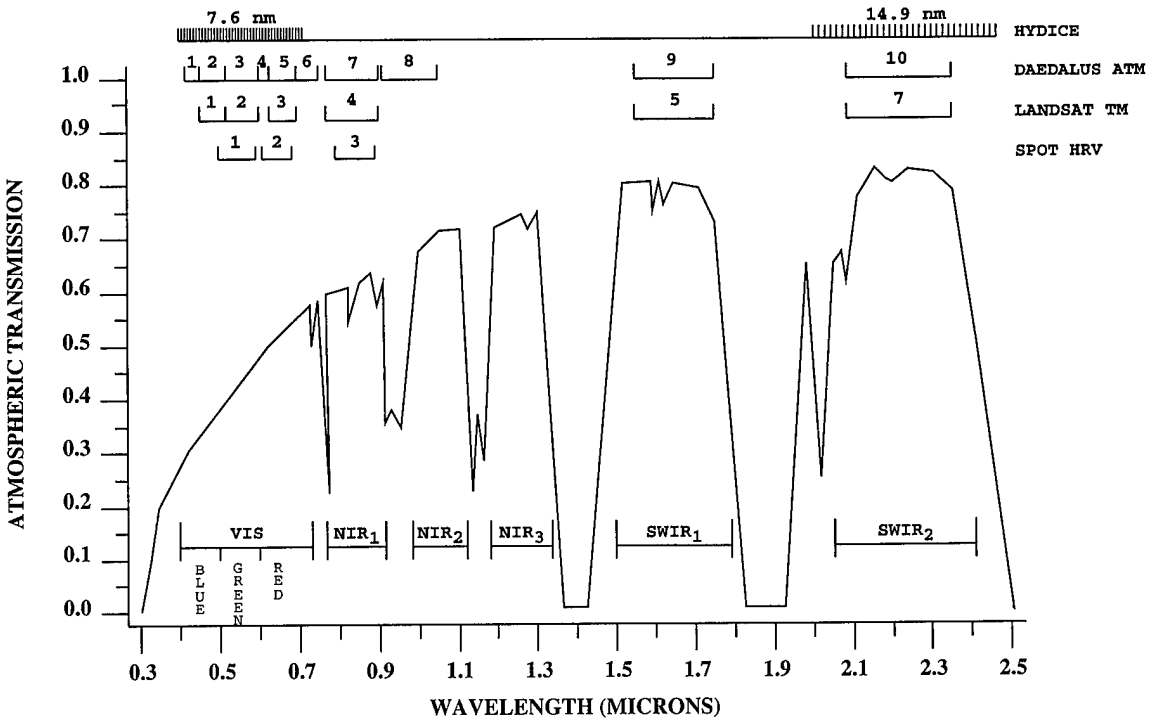


Figure 30. HYDICE sensor spectral bandpasses.

Ancillary navigation and environmental information is also recorded during the acquisition of HYDICE imagery. This information includes Inertial Navigation System (INS) data, Global Positioning System (GPS) data, flight stabilization platform data, and engineering data with instrument temperatures and voltages.

8.2 Fort Hood Data Collection

The collection of data at Fort Hood included both airborne imagery and ground truth measurements during 24–27 October 1995. Ground truth measurements are an essential part of the data collection process for characterization of surface materials to be imaged and atmospheric conditions at time of image acquisition. A significant amount of time and effort was spent in gathering ground truth data. In the following sections, the image acquisition and ground truthing activities conducted at Fort Hood are described.

8.2.1 Hyperspectral Image Collection

The image acquisition comprised of hyperspectral imagery collected by the HYDICE sensor system and natural color film shot by a KS–87 reconnaissance camera mounted on a CV–580 aircraft operated by Environmental Research Institute of Michigan (ERIM) of Ann Arbor, MI. Figure 31 shows the nine planned HYDICE flight-lines over Fort Hood's motor pool and barrack areas along with location of the six-step gray scale panel at Robert Gray Army Airfield projected on the subimage of a SPOT HRV multispectral (XS) scene. Each HYDICE flight-line has a ground sample distance (GSD) of 2 meters from a height of approximately 13,000 feet above ground level with 150 meter overlap between successive flight-lines to compensate for aircraft navigational errors. The resulting flight-lines possess a 640 meter cross-track and

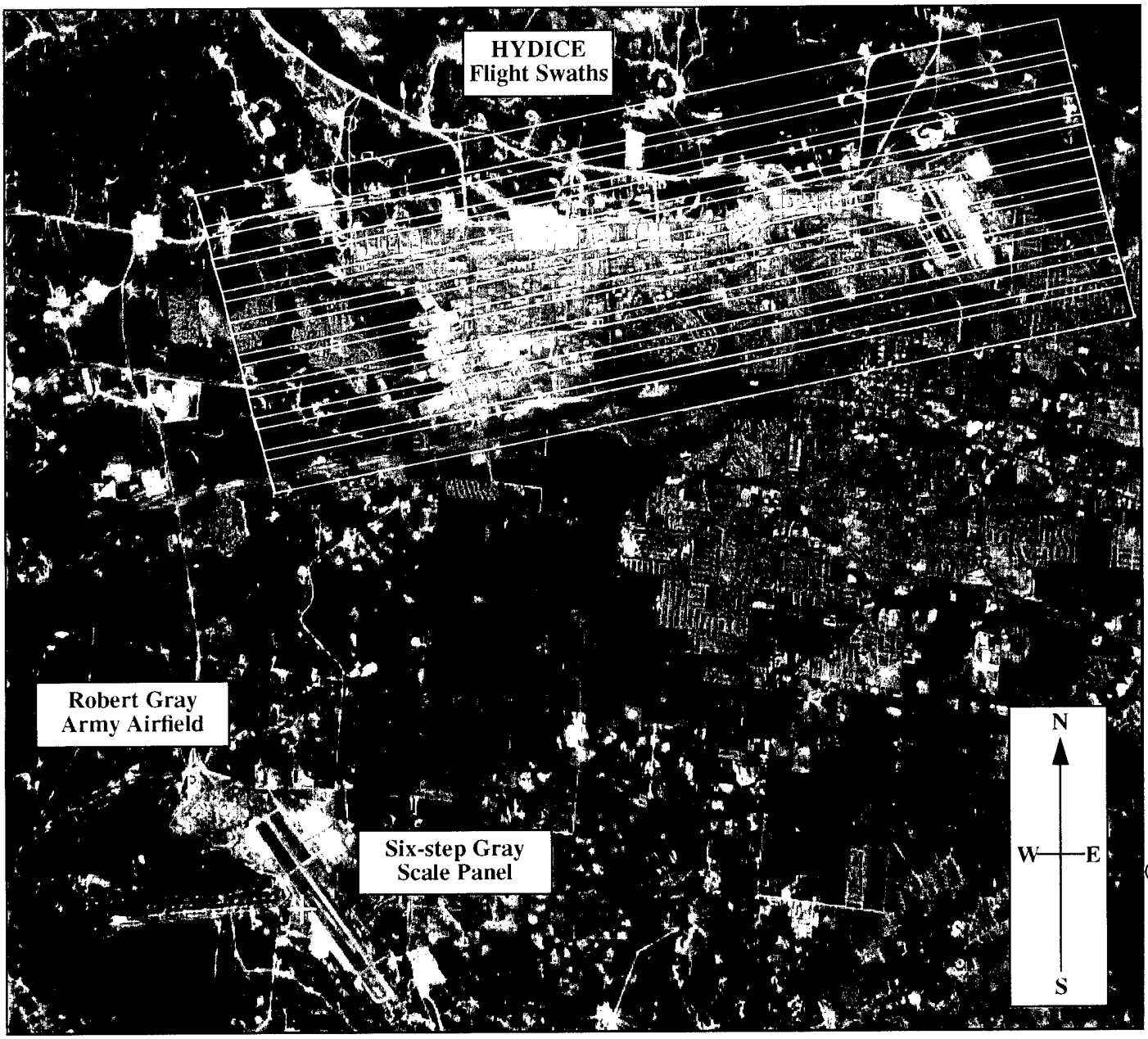


Figure 31. Fort Hood HYDICE flight coverage on FHSPOT1XS (SPOT HRV XS).

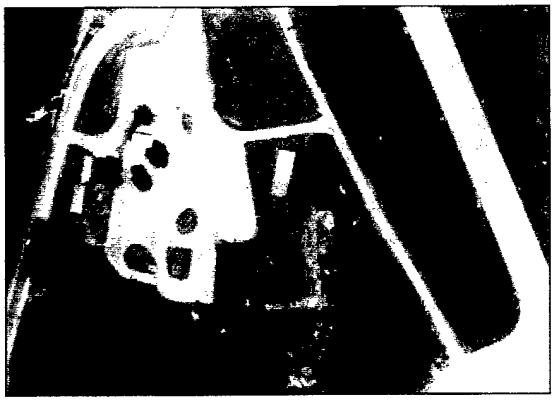


Figure 32. Aerial view of six-step gray scale panel.

Figure 33. Ground view of six-step gray scale panel.

12.6 kilometer along-track swath. A differential GPS station was established at Temple, TX airport to provide positioning information of the HYDICE sensor during image acquisition.

The flight-lines were flown in an east-to-west direction, beginning with the northernmost flight-line and proceeding in a southerly progression. During the return flight to the eastern edge of the next flight-line run, the HYDICE sensor was flown over and imaged the six-step gray scale panel at Robert Gray Army Airfield, providing crucial in-scene radiometric calibration measurements for each flight-line. Figures 32 and 33 show the deployment of the gray scale panel as captured by aerial reconnaissance and hand-held cameras, respectively. The six-step gray scale panel measured 30 by 180 feet with gray levels of 2, 4, 8, 16, 32 and 64 percent reflectance.

Normally, the six-step gray scale panel would be contained within each flight-line for in-scene radiometric calibration measurements. Because of the logistics of packing, transporting and deploying the gray scale panel for each flight-line within the time constraints (approximately a 15 minute window) set by image acquisition conditions, it was not feasible to have the gray scale panel located within each flight-line. To circumvent this logistic problem, a "race-track" flight path was used as previously described to achieve in-scene radiometric calibration measurements of the gray scale panel for each flight-line. The main assumption made assumed the atmospheric conditions over each flight-line and the gray scale panel were nearly the same. To monitor the atmospheric conditions, downwelling radiance measurements were collected at pre-determined sites contained in each flight-line as part of the ground truthing. The distance from these downwelling radiance measurement sites to the gray scale panel varied from 11.1 to 7.7 kilometers.

Originally, it was planned to collect the HYDICE imagery on 25 October 25, but this was cancelled becaluse of rain and thunderstorms. A second attempt took place on 26 October but was aborted due to Army Science Board exercises and 80 percent cloud cover. On the last possible day, 27 October, HYDICE imagery was collected after morning clouds gave way to clear afternoon skies. The first flight-line (northernmost) was re-acquired in the afternoon because of cloudy conditions in the morning.

8.2.2 Ground Truth Collection

In support of ground truth activities, MTL Systems of Dayton, OH collected both spectral radiometric and meteorological data to characterize the in-scene environment at Fort Hood, TX. This scene characterization data set includes measurements made during the HYDICE image acquisition flights as well as selected man-

made and natural background features located within the HYDICE flight swaths in Figure 31. The ground truth data set collection includes:

- Spectral radiometric data:
  - background reflectance.
  - six-step gray scale panel reflectance.
  - downwelling radiance in each HYDICE flight-line.
- Meteorological data:
  - surface weather parameters.
  - upper atmosphere parameters.

Background reflectance measurements of various man-made and natural materials were measured with a Geophysical Environmental Research (GER) dual beam spectral radiometer covering the 357 to 2500 nanometer region. Depending on the material being measured, the spectral radiometer was mounted on a wheeled-tripod or a 50-foot boom truck. Accessibility and solar illumination dictated which surface materials were measured. Accessibility issues ranged from the inability to position the spectral radiometer over an object because of limited maneuverability of the boom truck to restricted area access in barrack compounds. To obtain a reliable surface material spectral reflectance, direct solar illumination of the object was necessary to maintain a high signal-to-noise ratio with the spectral radiometer, especially in the near to shortwave infrared regions. The following list highlights the materials measured for spectral reflectance in the Fort Hood complex area:

- Man-made surface materials:
  - asphalt
  - concrete
  - bare and painted sheet metal roofing
  - desert camouflaged vehicle
- Natural surface materials:
  - clay soil/grass areas
  - deciduous (live oak) and coniferous (juniper) tree
  - grassland/rangeland
  - gravel.

Figures 34 and 35 show the boom truck being used to field measure a coniferous (juniper) tree and a desert camouflaged vehicle in a public-accessible parking lot. When possible, several different surface materials were measured during a boom truck setup. For example, grassland areas near the coniferous were collected while asphalt areas in the parking lot and concrete sidewalks next to the vehicles were measured.

To provide in-scene radiometric calibration for each HYDICE flight-line, the spectral reflectance of the six-step gray scale panel located at Robert Gray Army Airfield was measured prior to the start of the HYDICE flight collection and the downwelling radiance was collected in each HYDICE flight-line. At

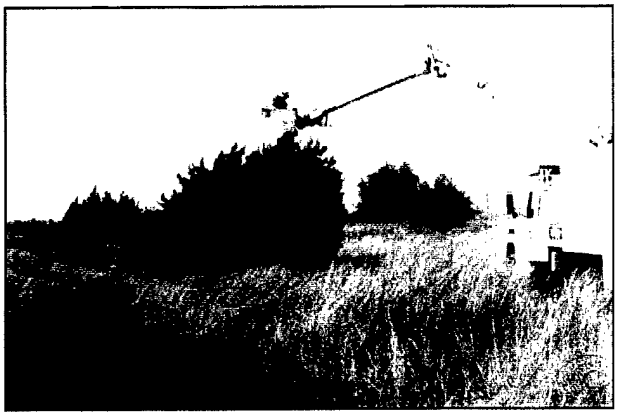


Figure 34. Coniferous tree reflectance measurement.

Figure 35. Desert camouflaged vehicle reflectance measurement.

pre-determined sites, the downwelling radiance (direct solar and sky) was measured by the GER spectral radiometer as the HYDICE sensor passed over the site. By making the assumption that the atmospheric conditions are nearly the same at the gray scale panel location, an in-scene calibration can be established for each flight-line.

Both surface and atmospheric meteorological data was to be collected during the HYDICE over-flight period. Surface weather parameters measured by an automated weather station located near the six-step gray scale panel included:

- ambient air temperature
- barometric pressure
- wind speed and direction
- relative humidity
- precipitation
- solar irradiance.

The solar irradiance data consisted of both direct solar and sky irradiance, as well as shadow band data (sky irradiance). Atmospheric data were to be derived from a weather balloon with an attached radiosonde system that measures air temperature, barometric pressure and relative humidity.

Surface weather parameters were measured on 25 and 26 October by the automated weather station. Upper atmospheric parameters were collected on 26 October at nearby Robert Gray Army Airfield to an altitude of 5000 feet before losing radiosonde transmission because of aircraft communication traffic. Both systems were unavailable on 27 October during the HYDICE over-flights since they were being transported to California. However, surface weather parameters for 27 October were supplied by the 3rd Weather Squadron at Robert Gray Army Airfield during the HYDICE over-flights.

8.3 HYDICE data status and planned processing

As of this writing, the frame reconnaissance imagery obtained on the flight has been developed and printed and the ground truth report is being finished. Once this is available, the HYDICE imagery can be calibrated to convert the digital counts recorded for each band into actual spectral radiance values.

The first part of the processing plan is to deal with the quantities of HYDICE image data, once it arrives. Preliminary estimates are about 4 gigabytes of data; calibration and post-processing may increase this.

Spectral processing, after calibration has been performed, will first involve merging channels into some smaller representation. Merging will be done using radiance measures, instead of just the raw digital counts, so that factors such as atmospheric transmission can be properly accounted for. Once a representative set of channels has been produced, classification efforts will begin.

In order to register the HYDICE data, the existing mapping imagery over Ft. Hood will be used as a baseline to select control and tie points visible in the HYDICE images. A simultaneous adjustment between the HYDICE and frame imagery will be performed to obtain the best possible registration. Geometric constraints, such as constraining roads to be straight lines [McGlone and Mikhail, 1982], also will be included to compensate for the dynamic nature of the HYDICE sensor.

9 Simulation

Constructing large-scale virtual world databases for ground-based simulation requires the integration of information from various sources, including digital map data, aerial and satellite imagery, detailed line drawings, and ground-based photography. Such virtual world databases have significant applications in DoD training, mission planning and rehearsal, and autonomous agent simulation. Our early work in this area primarily involved the construction of Triangular Irregular Networks (TINs) that formed a simple but efficient bare earth terrain skin [Polis and McKeown, 1993]. The issue was to intelligently select a small subset of the elevation points available in a digital elevation model (DEM) for inclusion in the TIN. The ability to reduce points by over two orders of magnitude was demonstrated while maintaining a high fidelity terrain representation. This permits real time graphics rendering using a modest polygon count.

At that time detailed man-made and natural features could be situated, usually manually, with heavy reliance on image and object textures to give the appearance of a detailed environment. Our research evolved into experiments with the direct integration of man-made features, initially roads, directly into the terrain skin. This allowed for a more realistic visualization and, more importantly, permitted roads to be automatically modeled to obey physical constraints with respect to road grade and side slope [Polis *et al.*, 1995]. Road trafficability became an issue within ground simulation both in the context of manned simulators and for computer generated forces. In the case of manned simulators, severe slopes on non-integrated roads caused the underlying physical mobility models to detect violations in trafficability even though the driver was traveling on a well defined road. Computer generated forces using terrain slope as a constraint for path planning might have to avoid perfectly trafficable areas or follow roads without regard to their geometry.

Our more recent research has focused on the integration of increasingly more complex man-made and natural features. Each feature generally brings a different set of physical constraint requirements into the virtual world generation process. Again, these may be dictated by the visual simulation requirements; that the virtual world computer graphics "look good," or by requirements placed on the representation by computer generated autonomous agents that need a consistent 3-D world for planning, navigation, and execution of intelligent behaviors.

Figure 36. Lake integrated into rugged terrain.

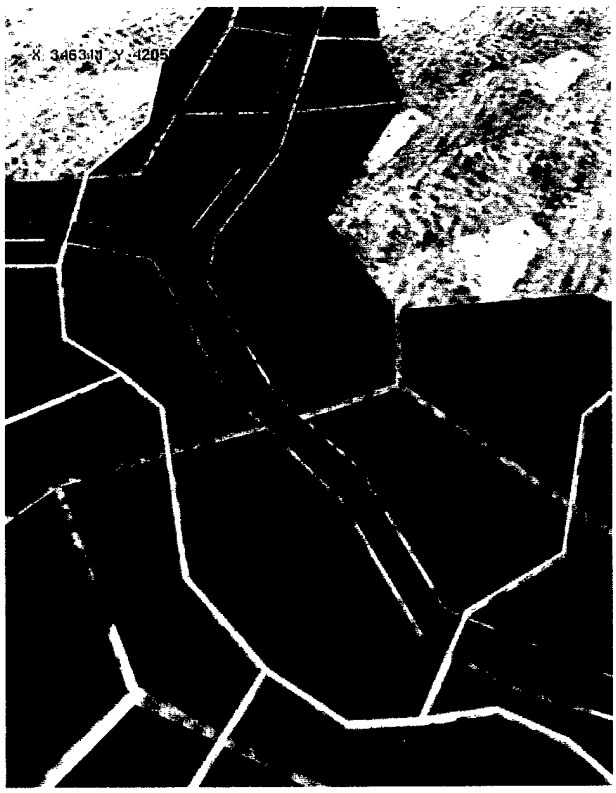


Figure 37. Rice paddies in river valley.

For example, from a visual standpoint, lakes should lie in a horizontal plane whose elevation is determined from the DEM such that it interacts in a natural way with the surrounding terrain. That is, generally lakes have smoothly sloping boundaries, not usually characterized as cliffs or similar sharp discontinuities with the terrain. However, when we are generating a TIN surface, some areas of the terrain surrounding the lake may not be selected based upon criteria derived from minimizing error with respect to the bare earth DEM surface. Therefore it is necessary to integrate the lake outline first, possibly simplifying the perimeter, and then allowing point selection to generate the surrounding terrain shape. Even in these simple cases, there can be mismatches between the DEM and the geo-position of the cartographic features. Such mismatches require more sophisticated reasoning regarding either displacement of the feature, simplification of the feature boundary, modification to the DEM, or some combination of these techniques.

However, if a simple polygonal feature such as a lake, constrained to a horizontal plane, potentially requires such a myriad of geometric analysis, consider constraints on more complicated features such as rivers, tree canopies, bridges, and railroads, complex composite structures such as rice paddies or cloverleaf overpasses, as well as interactions between individual features such as bridges and rivers, roads and bridges, etc.

Figures 36 and 37 show a small portion of the Prairie Warrior (Chorwan) virtual world database. Figure 36 shows a lake nestled in a complex mountainous area where the terrain appears to naturally descend into the lake boundary. Figure 37 shows the interaction between rice paddies in a river valley, with dike structures, roads, and a river running through the rice paddy structure. In both cases these features were integrated into the simulation terrain skin fully automatically, requiring minimal manual adjustments and intervention after examination of the resulting integrated TIN. We are excited by the opportunity to

demonstrate the integration of cartographic data while addressing accurate geo-positioning and coupled with results of automated feature extraction from imagery to populate these simulation databases. A more detailed set of examples showing our progress in automated virtual world construction can be found in these proceedings [McKeown *et al.*, 1996].

10 Dataset Acquisition, Processing, and Interchange

While nearly all research projects allude to the necessity of having adequate data sets for development and testing, very few actually make the effort to acquire and process such data sets. Indeed, of the most useful side effects of the RADIUS project has been the acquisition and distribution of large image datasets, including two sets of modelboard images [Thornton and others, 1994] and two sets of mapping images taken over Ft. Hood, Texas.

We have expended significant efforts in the last two years to obtain and process meaningful data sets. These have included blocks of aerial imagery for the development and evaluation of feature extraction algorithms (including the RADIUS Ft. Hood imagery), multispectral and hyperspectral data (Section 8) for experiments in high-resolution land cover mapping and data fusion, and digital feature and elevation data for simulation database generation (Section 9, [McKeown *et al.*, 1996]). While this involves major costs in money, time, and storage, we feel strongly that rigorous algorithm development and evaluation cannot be done on the toy data sets, disconnected from any real-world information, that have been the standard in the computer vision field.

Data set issues are not limited to imagery; our work on the production of simulation databases has given us painful experience in the problems of merging existing digital data of different sources. Both geometric issues, such as differing datums, and semantics issues, such as attribution, must be addressed if the merged dataset is to be more than a pretty display.

In order to deal with large volumes of data we are developing tools to aid in processing incoming image sets (Idl_Landmark) and to visualize and interrogate the various types of datasets, both image and feature, existing within our laboratory (Idl_Concept).

A related problem is how to exchange such datasets with other sites, using different systems. We have been active in promoting standard sensor model interchange formats within the RADIUS community and are now working on definitions of site model interchange formats, as described below.

10.1 Idl_Landmark—Dataset acquisition and processing

Idl_Landmark is a utility for viewing, editing and maintaining the landmark database. Landmarks are points used for image orientation and registration, with geo-coordinates and a text description. In addition to editing existing landmarks, Idl_Landmark provides an interface for performing control and tie point measurements, allowing simultaneous entry in multiple images. Measuring landmarks in new images is facilitated by Idl_Landmark's ability to project hypothetical landmark locations into the new image, reducing the operator's search time.

Plans include the addition of a digitizer interface for direct entry of coordinates from topographic maps, as well as full integration with the resection software.

10.2 Idl_Concept—Dataset access

Idl_Concept is an interactive system for displaying and accessing information about images and digital data sets, based on the ConceptMap system [McKeown, 1987].

Images are organized by geographic location, providing groups of images that a user would be likely to work with. General information on each image, such as image location and orientation information, is readily available. SceneDB, our database of standard subimages used for testing, also is supported, allowing the user to display and query scenes as well as add scenes to the database.

Several photogrammetric tools and functions also are provided for evaluation purposes. Image points or regions can be queried for coverage in other generic images, allowing the user to display and examine the region from other viewpoints. Idl_Concept also will display data coverage over a particular image or region, providing information on what ITD, DEM, or site data covers that area. Cursor tracking is available to perform active image to image projections.

Plans include providing more context-based image queries, such as queries of particular features, and using Idl_Concept as a front end for database preparation and selection.

10.3 Site model interchange

The importance of site model interchange cannot be understated; it allows different research organizations to share results, and to provide a common format for users of site models. However, it is difficult to develop a good interchange format because it needs "*to be usable by a wide variety of systems implemented in different languages and developed on diverse architectures*"[1]. Interchange of site models for image understanding and cartographic applications requires more than the simple output of object shape and size. In particular, it is important to store lineage information about the exchanged objects. This information allows others to understand how the objects were created, to understand the accuracy and precision of these objects, and to use this information to enhance the site model. The minimum requirements for the lineage information are the interchange of sensor models and the storage of image measurements that were computed by the system or measured by users. Furthermore, if the site model is to be used for cartographic applications such as the construction of a simulation database, exchanged objects must be stored in a geodetic coordinate system, or a local cartesian system tied to the geodetic system, with covariance information associated with each object. This information is required if the interchanged site models are to be merged with other type of geodetic information, such as DTED or ITD.

We have developed an interchange format and an application programmers interface(API)[2] that addresses these issues. The CMU MAPSLab Site Exchange Format (MASEF) has the ability to specify the camera/sensor model for single or multiple images by using existing RADIUS standard interchange formats: TEC header files and FBIP interchange files. Several types of building models of different complexities are defined, ranging from simple specific models, such as a rectangular flat roof building model, to more complex generic building objects. The relationship of all vertices in the building object to the topology of the model is explicitly defined. This relationship allows us to associate image measurements with model points for all images from which the model was constructed, and to store the covariance information for each model point.

---

[1] *IUE Data Exchange*, July 1995, p. 1-1.

[2] This is available at ftp.cs.cmu.edu. Login as `anonymous` and change your working directory to `/afs/cs/project/vdata-86/ftp`.

References

Barnard, 1983. S. Barnard. Interpreting perspective images. *Artificial Intelligence*, 21:435–462, 1983.

Biederman, 1985. I. Biederman. Human image understanding: Recent research and a theory. *CVGIP*, 32:29–73, 1985.

Cochran, 1994a. S. D. Cochran. Adaptive vergence for stereo matching. In *Int. Archives of Photogrammetry and Remote Sensing: Spatial Information from Digital Photogrammetry and Computer Vision*, vol. 30, 3/1, pp. 144–151, Munich, Germany, 5–9 Sept. 1994. Commission III.

Cochran, 1994b. S. D. Cochran. Adaptive vergence for the stereo matching of oblique imagery. In *ARPA IUW*, pp. 1335–1348, Monterey, CA, 13–16 Nov. 1994. ARPA, Morgan Kaufmann.

Cochran, 1995. S. D. Cochran. Adaptive vergence for the stereo matching of oblique imagery. *PERS*, 50(4):21–28, July/Aug. 1995.

Ford and McKeown, 1992a. S. J. Ford and D. M. McKeown, Jr. Utilization of multispectral imagery for cartographic feature extraction. In *DARPA IUW*, pp. 805–820, San Diego, CA, Jan. 1992. DARPA, Morgan Kaufmann.

Ford and McKeown, 1992b. S. J. Ford and D. M. McKeown, Jr. Information fusion of multispectral imagery for cartographic feature extraction. In *Int. Archives of Photogrammetry and Remote Sensing: Interpretation of Photographic and Remote Sensing Data*, vol. XVII, B7, DC, 2–14 Aug. 1992. XVIIth Congress, Commission VII.

Ford and McKeown, 1993. S. J. Ford and D. M. McKeown, Jr. Information fusion of multispectral imagery for cartographic feature extraction. In D. G. Leckie and M. D. Gillis, editors, *Proceedings of the International Forum on Airborne Multispectral Scanning for Forestry and Mapping (with Emphasis on MEIS)*, Information Report (Petawawa National Forestry Institute) PI–X–113, pp. 129–143, Val-Morin, Quebec, Canada, 13–16 Apr. 1993. Forestry Canada, Petawawa National Forestry Institute.

Ford *et al.*, 1993. S. J. Ford, J. B. Hampshire and D. M. McKeown, Jr. Performance evaluation of multispectral analysis for surface material classification. In *DARPA IUW*, pp. 421–435, DC, Apr. 1993. DARPA, Morgan Kaufmann.

Förstner, 1987. W. Förstner. Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision. *CVGIP*, 40:273–310, 1987.

Förstner, 1994. W. Förstner. Diagnostics and performance evaluation in computer vision. In *Proceedings, NSF/ARPA Workshop on Performance versus Methodology in Computer Vision*, pp. 11–25, 1994.

Fritsch *et al.*, 1994. D. Fritsch, M. Sester and T. Schenk. Test on image understanding. In *Proceedings: ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision, Volume 30, Part 3/1*, pp. 243–248, Munich, Germany, 1994.

Fua and Hanson, 1991. P. Fua and A. J. Hanson. An optimization framework for feature extraction. *Machine Vision and Applications*, 4(2):59–87, 1991.

Gee and Newman, 1993. S. Gee and A. Newman. RADIUS: Automating image analysis through model-supported exploitation. In *DARPA IUW*, pp. 185–196, DC, 19–21 Apr. 1993. Morgan Kaufmann Publishers, Inc.

Hsieh, 1995. Y. Hsieh. Design and evaluation of a semi-automated site modeling system. Technical Report CMU–CS–95–195, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, Nov. 1995.

Hsieh, 1996. Y. Hsieh. Design and evaluation of a semi-automated site modeling system. In *ARPA IUW*, Palm Springs, California, Feb. 1996. ARPA, Morgan Kaufmann.

Huertas and Nevatia, 1988. A. Huertas and R. Nevatia. Detecting buildings in aerial images. *CVGIP*, 41(2):131–152, Feb. 1988.

Irvin and McKeown, 1989. R. B. Irvin and D. M. McKeown. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. *IEEE Trans. SMC*, 19(6):1564–1575, 1989. Also available as Technical Report CMU–CS–88–200, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213.

Jaynes *et al.*, 1994. C. Jaynes, F. Stolle and R. Collins. Task driven perceptual organization for extraction of rooftop polygons. In *Proceedings: ARPA Image Understanding Workshop*, vol. I, pp. 359–365, Nov. 1994.

Lin *et al.*, 1994. C. Lin, A. Huertas and R. Nevatia. Detection of buildings using perceptual grouping and shadows. In *CVPR*, pp. 62–69, Seattle, WA, 19–23 June 1994.

Liow and Pavlidis, 1990. Y.-T. Liow and T. Pavlidis. Use of shadows for extracting buildings in aerial images. *CVGIP*, 49(2):242–277, Feb. 1990.

Lowe, 1985. D. G. Lowe. *Perceptual Organization and Visual Recognition.* Kluwer, Boston, 1985.

McGlone and Mikhail, 1982. C. McGlone and E. M. Mikhail. Geometric constraints in multispectral scanner data. In *Proceedings of the 48th Annual Meeting, American Society of Photogrammetry*, pp. 563–572. American Society of Photogrammetry, Mar. 1982.

McGlone and Shufelt, 1993. J. C. McGlone and J. A. Shufelt. Incorporating vanishing point geometry into a building extraction system. In *DARPA IUW*, pp. 437–448, DC, 19–21 Apr. 1993. DARPA, Morgan Kaufmann.

McGlone and Shufelt, 1994a. J. C. McGlone and J. A. Shufelt. Projective and object space geometry for monocular building extraction. In *CVPR*, pp. 54–61, Seattle, WA, 19–23 June 1994.

McGlone and Shufelt, 1994b. J. C. McGlone and J. A. Shufelt. Projective and object space geometry for monocular building extraction. Technical Report CMU–CS–94–118, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, 1994.

McGlone, 1992. J. C. McGlone. Design and implementation of an object-oriented photogrammetric toolkit. In *Int. Archives of Photogrammetry and Remote Sensing*, vol. XXIX, B2, pp. 334–338, 1992.

McGlone, 1995. C. McGlone. Bundle adjustment with object space constraints for site modeling. In *Proc. SPIE: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision*, vol. 2486, pp. 25–36, Apr. 1995.

McKeown and Denlinger, 1988. D. M. McKeown and J. L. Denlinger. Cooperative methods for road tracking in aerial imagery. In *CVPR*, pp. 662–672, Ann Arbor, MI, June 1988.

McKeown and McGlone, 1993. D. McKeown and J. C. McGlone. Integration of photogrammetric cues into cartographic feature extraction. In *Proc. SPIE: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision*, vol. 1944, pp. 2–15, Sept. 1993.

McKeown and Perlant, 1992. D. McKeown and F. Perlant. Refinement of disparity estimates through the fusion of monocular image segmentations. In *CVPR*, pp. 486–492, Urbana-Champaign, IL, 16-18 June 1992.

McKeown *et al.*, 1994. D. M. McKeown, Jr. *et al.* Research in automated analysis of remotely sensed imagery: 1993–1994. In *ARPA IUW*, pp. 99–132, Monterey, CA, 13–16 Nov. 1994. ARPA, Morgan Kaufmann. Also available as Technical Report CMU–CS–94–197, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213.

McKeown *et al.*, 1996. D. M. McKeown *et al.* Progress in the automated construction of virtual worlds. In *ARPA IUW*, Palm Springs, California, Feb. 1996. ARPA, Morgan Kaufmann.

McKeown, 1987. D. M. McKeown. The role of artificial intelligence in the integration of remotely sensed data with geographic information systems. *IEEE Trans. GRS*, GE–25(3):330–348, May 1987. Also available as Technical Report CMU–CS–86–174 School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213.

McKeown, 1990. D. M. McKeown. Toward automatic cartographic feature extraction. In L. F. Pau, editor, *Mapping and Spatial Modelling for Navigation*, vol. F 65 of *NATO ASI Series*, pp. 149–180. Springer-Verlag, Berlin Heidelberg, 1990.

McKeown, 1996. D. M. McKeown, Jr. Top ten lessons learned in automated cartography. Technical Report CMU–CS–96–110, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, Jan. 1996.

Mikhail, 1970. E. M. Mikhail. Relative control for extraterrestrial work. *Photogrammetric Engineering*, 36(4), Apr. 1970.

Mikhail, 1980. E. M. Mikhail. *Observations and Least Squares.* Harper and Row, New York, 1980.

Mohan and Nevatia, 1989. R. Mohan and R. Nevatia. Using perceptual organization to extract 3-D structures. *IEEE Trans. PAMI*, 11(11):1121–1139, Nov. 1989.

Nicolin and Gabler, 1987. B. Nicolin and R. Gabler. A knowledge-based system for the analysis of aerial images. *IEEE Trans. GRS*, GE–25(3):317–329, May 1987.

Norvelle, 1981. F. R. Norvelle. Interactive digital correlation techniques for automatic compilation of elevation data. Technical Report No. ETL-0272, U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA 22060, Oct. 1981.

Norvelle, 1992. F. R. Norvelle. Stereo correlation: Window shaping and DEM corrections. *PERS*, 58(1):111–115, Jan. 1992.

Pentland, 1986. A. P. Pentland. Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28:293–331, 1986.

Polis and McKeown, 1993. M. F. Polis and D. M. McKeown, Jr. Issues in iterative TIN generation to support large scale simulations. In AUTOCARTO 11*: International Symposium on Computer Assisted Cartography*, pp. 267–277, Minneapolis, MN, 30 Oct.–1 Nov. 1993.

Polis *et al.*, 1995. M. F. Polis, S. J. Gifford and D. M. McKeown, Jr. Automating the construction of large scale virtual worlds. *IEEE Computer*, 28(7):57–65, July 1995.

Pope, 1975. A. J. Pope. The statistics of residuals and the detection of outliers. Technical Report NOS 65 NGS 1, National Oceanic and Atmospheric Administration, 1975.

Roux and McKeown, 1994a. M. Roux and D. M. McKeown, Jr. Feature matching for building extraction from multiple views. In *CVPR*, pp. 46–53, Seattle, WA, 19–23 June 1994.

Roux and McKeown, 1994b. M. Roux and D. M. McKeown, Jr. Feature matching for building extraction from multiple views. In *ARPA IUW*, pp. 331–349, Monterey, CA, 13–16 Nov. 1994. ARPA, Morgan Kaufmann.

Roux *et al.*, 1995. M. Roux, Y. C. Hsieh and D. M. McKeown, Jr. Performance analysis of object space matching for building extraction using several images. In *Proc. SPIE: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II*, vol. 2486, pp. 277–297, 1995.

Shufelt and McKeown, 1993. J. A. Shufelt and D. M. McKeown. Fusion of monocular cues to detect man-made structures in aerial imagery. *CVGIP*, 57(3):307–330, May 1993.

Shufelt, 1996. J. Shufelt. Performance evaluation and analysis of vanishing point detection techniques. In *ARPA IUW*, pp. 1113–1132, Palm Springs, California, Feb. 1996. ARPA, Morgan Kaufmann.

Strunz, 1992. G. Strunz. Feature based image orientation and object reconstruction. In *International Archives of Photogrammetry and Remote Sensing*, vol. XXIX, B3, pp. 113–118, 1992.

Thornton and others, 1994. K. Thornton *et al.* Groundtruthing the RADIUS model-board imagery. In *ARPA IUW*, pp. 319–329, Nov. 1994.

Zlotnick and Carnine, 1993. A. Zlotnick and P. D. Carnine. Finding road seeds in aerial images. *CVGIP*, 57(2):243–260, Mar. 1993.